Genomic Approaches to Dissect Innate Immune Pathways

Dissertation Advisor: Dr. Nir Hacohen Mark N. Lee

Genomic approaches to dissect innate immune pathways

Abstract

The innate immune system is of central importance to the early containment of infection. When receptors of innate immunity recognize molecular patterns on pathogens, they initiate an immediate immune response by inducing the expression of cytokines and other host defense genes. Altered expression or function of the receptors, the molecules that mediate the signal transduction cascade, or the cytokines themselves can predispose individuals to infectious or autoimmune diseases. Here we used genomic approaches to uncover novel components underlying the innate immune response to cytosolic DNA and to characterize variation in the innate immune responses of human dendritic cells to bacterial and viral ligands.

In order to identify novel genes involved in the cytosolic DNA sensing pathway, we first identified candidate proteins that interact with known signaling molecules or with dsDNA in the cytoplasm. We then knocked down 809 proteomic, genomic, or domain-based candidates in a high-throughput siRNA screen and measured cytokine production after DNA stimulation. We identified ABCF1 as a critical protein that associates with DNA and the known DNA-sensing components, HMGB2 and IFI16. We also found that CDC37 regulates stability of the signaling molecule, TBK1, and that chemical inhibition of CDC37 as well as of several other pathway regulators (HSP90, PPP6C, PTPN1, and TBK1) potently modulates the innate immune response to DNA and to retroviral infection. These proteins represent potential therapeutics targets for infectious and autoimmune diseases that are associated with the cytosolic DNA response.

We also developed a high-throughput functional assay to assess variation in responses of human monocyte-derived dendritic cells to LPS (receptor: TLR4) or influenza (receptors: RIG-I and TLR3), with the goal to ultimately map genetic variants that influence expression levels of pathogen-responsive genes. We compared the variation in expression between the dendritic cells of 30 different individuals, and within paired samples from 9 of these individuals collected several months later. We found genes that have significant inter- vs. intra-individual variation in response to the stimuli, suggesting that there is a substantial genetic component underlying variation in these responses. Such variants may help to explain differences between individuals' risk for infectious, autoimmune, or other inflammatory diseases.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

ACKNOWLEDGMENTS

I would like to first acknowledge my advisor, Dr. Nir Hacohen, for all his support and guidance while I have been in the lab, and for sharing with me his passion for research. I have learned a ton from discussing projects with Nir, going over grants, and talking about larger ideas in science and medicine. It has been fantastic to work with someone who has such a vision about the field. I owe you much gratitude.

I also owe much to my cubical mates at the Broad Institute, my home away from home. Chloe Villani, Marciela DeGrace, Sagi Shapira, Dan-Avi Landau, and our honorary cubical member Jimmie Ye – it has been a pleasure to share a space with you. You have kept me entertained and have been my confidantes and counselors in good times and in harder times. Thank you for your support and friendship.

I would like to thank my collaborators. To Chloe, who has been both a teacher and second mother to me (and to so many of the other students), I do not know how I could have done this without you. Thank you for your amazing generosity and patience. Thank you to Jimmie for always being there to talk about our research and the field; your enthusiasm is infectious. Thank you to Weibo Li for helping out in any way possible. I really appreciated working together with you. Thank you to Dr. William Pendergraft III for all the kindness you have shown me, for believing in the work and editing so diligently. Thank you to my collaborators, especially Shao-En Ong, Philipp Mertins, and Farokh Dotiwala, for all your work and help.

Thank you to my advisory committee, Dr. Nicholas Haining, Dr. Jonathan Kagan, Dr. Laurie Glimcher, and Dr. Michael Brenner. You have been not only extremely helpful during meetings, but also so supportive and I truly appreciate that.

Thank you to the other members of the Hacohen lab for all of your help. It has been a team effort to keep the lab running. All the best in the future to all of you.

INTRODUCTION

The innate immune system is of central importance to the early containment of infection. When receptors of innate immunity recognize molecular patterns on pathogens, they initiate an immediate immune response by inducing the expression of cytokines and other host defense genes. Altered expression or function of the receptors, the molecules that mediate the signal transduction cascade, or the cytokines themselves can predispose individuals to infectious or autoimmune diseases. Here we used genomic approaches to uncover novel components of the innate immune system and to characterize variation in innate immune responses between humans.

The detection of pathogens by the innate immune system

The traditional model of the innate immune system is that it detects pathogens by discriminating "self" from "non-self". In 1989, Charles Janeway wrote, "I contend that the immune system has evolved specifically to recognize and respond to infectious microorganisms, and that this involves recognition…of certain characteristic patterns common on infectious agents but absent from the host" (Janeway 1989). This model of host germline-encoded pattern recognition receptors (PRRs) detecting pathogen-associated molecular patterns (PAMP) was borne out by (i) identification in 1996 of the Drosophila melanogaster Toll gene which was found to control the synthesis of antimicrobial peptides and protect against widespread fungal infection in the insect (Lemaitre et al. 1996); and then (ii) identification in 1998 of human and mouse Toll-like receptor 4 (TLR4) which was found to be the PRR for the Gram negative bacterial cell wall component lipopolysaccharide (LPS) (Medzhitov et al. 1997; Poltorak et al. 1998).

This traditional model has been modified to include not only the recognition of pathogen-specific PAMPs but also the recognition of components shared between host and pathogen by recognizing these components in contexts specific to infection (e.g. DNA in the cytosol, discussed in further detail below); and the model has been expanded to include the recognition of host-derived signals of cellular stress (e.g. uric acid and ATP) referred to as danger-associated molecular patterns (DAMPs) (Schroder and Tschopp 2010).

Pattern recognition receptors (Schroder and Tschopp 2010; Takeuchi and Akira 2010) are expressed by immune cells – such as monocytes, dendritic cells, and macrophages – that form the first line of defense against infection, but also by certain cells of the adaptive immune system (such as B cells) and certain non-hematopoietic cells depending on PRR type. The PRRs include the membrane-bound TLRs that scan the extracellular and endosomal spaces, the RIG-like receptors (RLRs) that scan the cytosol for viral RNA, the NOD-like receptors (NLRs) that scan the cytosol for various microbial components and DAMPs, and DNA-sensing receptors including AIM2 and RNA polymerase III that scan the cytosol for viral and possibly bacterial DNA. Several PRR types that are highlighted in the subsequent chapters are discussed in more detail below.

Toll-like receptors (TLRs)

TLRs are transmembrane proteins consisting of N-terminal leucine-rich repeat (LRR) domains; transmembrane domains; and C-terminal Toll-interleukin 1 receptor (TIR) domains (Kawai and Akira 2010). The LRR domains mediate PAMP recognition, while the TIR domain initiates a signal transduction cascade downstream of the receptor. 10 functional TLRs have been identified in humans and 12 in mice. They recognize a diversity of microbial PAMPs, including microbial

lipids/lipoproteins (TLR1/2/4/6), proteins (TLR5/11), and nucleic acids (TLR3/7/8/9) derived from microorganisms including bacteria, viruses, parasites, and fungi (Akira and Takeda 2004). TLRs regulate the expression of pro-inflammatory genes including TNFα, IL-6, and pro-IL-1β; as well as type I and type III interferon (IFN) genes, which themselves control the expression of many anti-viral genes. Certain TLRs, e.g. TLR2, primarily regulate inflammatory gene induction, while other TLRs, e.g. TLR3, primarily regulate the induction of anti-viral genes (Amit et al. 2009).

Many of the TLRs are expressed on the cell surface, but the nucleic-acid sensing TLRs (TLR3, TLR7, TLR8, and TLR9) are expressed on the endoplasmic reticulum (ER) and then translocate to endolysosomes following ligand stimulation (Latz et al. 2004; Kim et al. 2008). The ER-resident protein, UNC93B1, regulates this translocation event (Tabeta et al. 2006; Kim et al. 2008). The cell types that express TLRs vary, with some TLRs specific to immune cells (e.g. TLR7 and TLR9 are only expressed in plasmacytoid dendritic cells and B cells in humans) and some more widely expressed (e.g. TLR3 is expressed in various hematopoietic as well as non-hematopoietic cells (Zhang et al. 2007)).

RIG-like receptors (RLRs)

Infection by RNA viruses in a broad range of cell types elicits inflammatory gene and IFN production mediated by the recognition of the virus's RNA genome in the cytosol by RIG-like receptors (Barbalat et al. 2011). There are three known RLRs – *RIG-I*, *Mda5*, and *Lgp2* – characterized by a C-terminal DEAD-box (Asp-Glu-Ala-Asp) helicase domain that mediates RNA binding and two N-terminal caspase recruitment domains (CARDs) that mediate downstream signaling (though *Lgp2* lacks the CARD domains) (Yoneyama et al. 2004). RIG-I is

believed to recognize 5' triphosphate groups on single-stranded RNA as well as short double-stranded RNA; MDA5 is believed to recognize long double-stranded RNA; and the function of LGP2 is less clear (it is believed to regulate the other RLRs, though data is conflicting) (Yoneyama et al. 2005; Satoh et al. 2010; Barbalat et al. 2011). The RLRs appear to be broadly expressed, as most hematopoietic and non-hematopoietic cell types tested respond to RLR ligand stimulation (with the possible exception of plasmacytoid dendritic cells (pDCs)) (Kato et al. 2005).

Cytosolic DNA sensors

Similar to the RLR response, cytosolic DNA has been shown to induce type I IFN production, though **the receptor(s) and components of this pathway are much less clear**. This phenotype was first described in 1962-1963 by several groups (Isaacs et al. 1963; Jensen et al. 1963; Rotem et al. 1963). They had suggested that DNA viruses such as vaccinia virus stimulate the production of IFN through viral genomic DNA, performed initial studies showing that adding viral or even mammalian DNA to cells elicits IFN production, and suggested that the DNA needs to accumulate within the cell to be stimulatory. In 1999, Suzuki e*t al*. demonstrated that transfection of herpes simplex viral DNA, mammalian DNA, or synthetic double-stranded DNA into a wide variety of cell types (thyrocytes, macrophages, dendritic cells, fibroblasts, and endothelial cells) increased expression of MHC class I and II (Suzuki et al. 1999). This effect was independent of sequence (e.g. the effect was not dependent on the presence of unmethylated CpG motifs which are recognized by TLR9). In 2006, Ishii *et al*. had shown that transfection of long stretches of double-stranded alternating AT base pairs robustly stimulates the production of type I IFN as well as NFκB-responsive genes in a manner independent of known TLR pathways

(Ishii et al. 2006). In the same year, Stetson *et al*. showed that transfected dsDNA of any sequence stimulates the production of type I IFNs in a manner independent of TLR9 (Stetson and Medzhitov 2006).

The pathway stimulated by Stetson *et al.* is referred to as the ISD (interferon stimulatory DNA) pathway (Stetson and Medzhitov 2006). While it is not yet entirely clear which pathogens (e.g. DNA viruses or intracellular bacteria) are sensed by the ISD pathway, perhaps the clearest data is that retroviral DNA (produced after reverse transcription of the genomic RNA) is sensed by this pathway (Yan et al. 2010). It is also not yet entirely clear what the receptor of the ISD pathway is. Several proteins that interact with DNA and bind to downstream signaling proteins have been proposed as receptors:

(i) DAI/ZBP1. In 2008, Takaoka *et al*. identified DAI (also known as ZBP1) as a cytosolic protein that recognizes transfected DNA, interacts with downstream signaling components (i.e. TBK1 and IRF3, discussed below), and regulates type I IFN production (Takaoka et al. 2007). However, cells from DAI-deficient mice did not confirm this phenotype, possibly suggesting redundancy and/or cell type specificity (Ishii et al. 2008). Evidence is emerging, though, that DAI can indeed activate NFκB and may be involved in the recognition of cytomegalovirus (Kaiser et al. 2008; Rebsamen et al. 2009; DeFilippis et al. 2010).

(ii) HMGB family proteins. In 2009, Yanai *et al.* identified the HMGB family proteins (HMGB2 and possibly HMGB3) as DNA interactors that regulate type I IFN production (Yanai et al. 2009). They found that all nucleic acid-sensing pathways of the innate immune system require HMGB family proteins to function, and suggested that these proteins act as co-receptors for nucleic acid recognition by the innate immune system.

(iii) LRRFIP1. In 2010, Yang *et al.* identified LRRFIP1 as a cytosolic protein that interacts with DNA and regulates type I IFN production (Yang et al. 2010). They suggested, though, that LRRFIP1 interacts with both DNA and RNA, and activates the co-activator, β-catenin, rather than known downstream components.

(iv) IFI16. In 2010, Unterholzner *et al.* identified IFI16 as another cytosolic protein that recognizes cytosolic DNA, interacts with a downstream signaling component (i.e. STING, discussed below), and regulates type I IFN production (Unterholzner et al. 2010). However, IFI16 was found to primarily localize to the nucleus, and knockdown of *Ifi16* only gave a partial (~2-fold) phenotype.

(v) KU70. In 2011, Zhang *et al.* identified KU70 as a DNA-interacting protein that regulates type III but not type I IFN production (Zhang et al. 2011a).

(vi) DDX41. In 2011, Zhang *et al.* identified DDX41 as a cytosolic protein that recognizes cytosolic DNA, interacts with a downstream signaling component (i.e. STING), and regulates type I IFN production (Zhang et al. 2011b). Their data suggested that DDX41 is active in dendritic cells, but a role for DDX41 in cells outside of the myeloid-lineage was less clear.

Thus, the identity of the receptor(s) of the ISD pathway is less clear than that of the other innate immune pathways. One possible reading of the literature is that DAI is not the ISD receptor because DAI-deficient cells respond normally to DNA stimulation; the HMGB proteins are co-receptors that aid receptor binding (perhaps through their ability to distort DNA) but are not sufficient as a receptor; LRRFIP1 helps regulate IFN through the alternative β-catenin arm but is also not sufficient; IFI16 recognizes foreign nuclear DNA in an alternative ISD pathway or helps regulate IFN but is also not sufficient; KU70 regulates type III but not type I IFN

production; and DDX41 is active in myeloid-lineage cells but not other cell types in which the ISD response has been observed. Such an interpretation would suggest the hypothesis that either the main ISD receptor has not yet been identified, or that the ISD receptor is a complex of two or more of the above proteins.

In addition, the cellular compartment in which DNA recognition occurs has also been debated in the literature recently (Unterholzner et al. 2010; Kerur et al. 2011; Li et al. 2012; Orzalli et al. 2012). Groups have suggested that DNA recognition may occur in the nucleus rather than the cytosol. This hypothesis was borne out of results showing that IFI16 is primarily localized to the nucleus, and that the DNA of certain viruses such as HSV-1 co-localize with IFI16 in the nucleus (Li et al. 2012). Such a hypothesis is not incompatible with previous results, given that DNA transfection using lipid reagents likely delivers DNA to most cellular compartments. The data that suggests a cytosolic receptor includes: the analogous RNA receptors are cytosolic (Yoneyama et al. 2004); the DNA receptor of the inflammasome, AIM2 (described below), is cytosolic (Burckstummer et al. 2009; Fernandes-Alnemri et al. 2009; Hornung et al. 2009); genomic DNA is nuclear and would have to be distinguished from foreign nuclear DNA; and the downstream components STING and TBK1 are cytosolic (Ishikawa and Barber 2008; Ishikawa et al. 2009) and so the signal would have to first exit the nucleus and then come back during IRF3 translocation. One possibility is that there are two different DNA receptors – one in the cytosol and one in the nucleus.

Following the publication by Ishii *et al.* (Ishii et al. 2006), it was subsequently shown that the AT-rich ligand used by their group not only stimulates the ISD pathway, but also stimulates the RNA polymerase III pathway. In this pathway, AT-rich DNA as well as DNA sequences with specific promoter elements are recognized by RNA polymerase III, which transcribes the

DNA and the resulting RNA product is then able to stimulate the RIG-I receptor (Ablasser et al. 2009; Chiu et al. 2009). While data suggests that certain viruses (e.g. HSV-1, Epstein-Barr virus, and adenovirus) and intracellular bacteria (e.g. *Legionella pneumophila*) are sensed by the RNA polymerase III pathway (Chiu et al. 2009), there is conflicting data in the literature (Unterholzner et al. 2010).

In addition to the ISD and RNA polymerase III pathways, an alternative pathway mediated by the AIM2 receptor (also known as the AIM-like receptor or ALR pathway) was shown to bind to DNA and post-translationally regulate IL-1β maturation (Muruve et al. 2008; Burckstummer et al. 2009; Fernandes-Alnemri et al. 2009; Hornung et al. 2009; Roberts et al. 2009). IFI202, which is similar in structure to AIM2 but lacks the C-terminal CARD domain, is believed to negatively regulate the AIM2 pathway (Roberts et al. 2009).

Signal transduction cascades in the innate immune system

Following engagement of PRRs by microbial ligands, the receptors initiate a signal transduction cascade that ultimately leads to transcriptional and post-translational changes in the cell. For the TLRs, RLRs, and certain NLRs and cytosolic DNA sensors, downstream signaling activates a common set of transcription factors including NFκB and AP-1 that induce the expression of inflammatory genes (e.g. IL-6, TNFα, IL-12); as well as IRF family members (e.g. IRF3 and IRF7) that induce the expression of type I and III IFNs (Akira and Takeda 2004; Barbalat et al. 2011). For AIM2 as well as certain NLR family members termed inflammasomes, receptor engagement triggers post-translational maturation of caspase-1, which regulates maturation and secretion of inflammatory cytokines such as IL-1β, IL-18, and IL-33 (Schroder and Tschopp 2010). The signal transduction cascades of several of these pathways are discussed in more detail

below. Each of these signaling cascades involves a multitude of proteins and interactions, and key signaling components are described.

TLR4. The intracellular TIR domain of TLR4 recruits the TIR-domain-containing adaptor MYD88, which controls the expression of inflammatory genes (Akira and Takeda 2004). TLR4 then translocates to endosomes where it interacts with the TIR-domain-containing adaptor TRIF, which controls the expression of both inflammatory and anti-viral genes (Kagan et al. 2008). The adaptor MYD88 recruits IL-1 receptor–associated kinases (IRAKs), which then activate the E3 ubiquitin ligase TNF receptor-associated factor 6 (TRAF6) to catalyze the synthesis of K63-linked polyubiquitin chains. The polyubiquitin chains then recruit the TAK1 kinase complex as well as the IkB kinase (IKK) complex member, NEMO (NFκB essential modulator), to activate the IKK catalytic kinase subunits (IKKα and/or IKKβ). These subunits phosphorylate IκB, an inhibitor that interacts with and traps NFκB in the cytosol. Phosphorylation of IκB (e.g. IκBα) initiates its proteasomal degradation, allowing NFκB to translocate into the nucleus and turn on the expression of target genes. The TAK1 kinase complex also activates MAP kinases (MAPKs), which activate the transcription factor, AP-1. In contrast to MYD88, the adaptor TRIF recruits the kinases TANK-binding kinase 1 (TBK1) and IKKε, which phosphorylate the transcription factor, IRF3 (Fitzgerald et al. 2003; McWhirter et al. 2004). Phosphorylated IRF3 dimerizes and translocates into the nucleus to induce the expression of IFN genes. In addition to TBK1 and IKKϵ activation, TRIF also recruits TRAF6 and the adaptor RIP1 to activate NFκB (Akira and Takeda 2004).

TLR3. Following recognition of double-stranded RNA (Alexopoulou et al. 2001), TLR3 recruits TRIF to activate downstream components as described above (Akira and Takeda 2004).

RIG-I. Upon binding to its RNA ligand, RIG-I is ubiquitinated by TRIM25 and undergoes a conformational change that exposes its CARD domains (Yoneyama et al. 2004; Gack et al. 2007). The CARD domains then form homotypic interactions with the CARD domain of the downstream adaptor, MAVS, on the mitochondrial surface and on peroxisomes (Kawai et al. 2005; Meylan et al. 2005; Seth et al. 2005; Dixit et al. 2010). Although the precise mechanism is controversial, MAVS ultimately activates NFκB and also recruits TBK1 to activate IRF3 (Kawai et al. 2005; Meylan et al. 2005; Seth et al. 2005).

ISD pathway. While the receptor RIG-I and the adaptor MAVS is not required for ISD signaling (Sun et al. 2006), the ISD pathway shares many of the downstream signaling components as the RIG-I pathway. The endoplasmic reticulum (ER)-membrane adaptor, STING (Ishikawa and Barber 2008), interacts with TBK1 and IRF3 resulting in IRF3 phosphorylation by TBK1 (Stetson and Medzhitov 2006; Ishikawa and Barber 2008; Ishikawa et al. 2009; Tanaka and Chen 2012).

Type I IFN receptor. Following type I IFN production, the IFNs are secreted and interact with a heterodimeric receptor composed of IFNAR1 and IFNAR2 (Uze et al. 1990; Novick et al. 1994). Upon IFN binding, the Janus kinases, JAK1 (Muller et al. 1993) and TYK2 (Velazquez et al. 1992), phosphorylate each other and the cytoplasmic domain of IFNAR1, allowing phosphorylation of STAT2 and then STAT1 (Fu et al. 1992; Schindler et al. 1992; Stark et al. 1998). The phosphorylated STAT proteins dimerize, interact with IRF9 to form the ISGF3 complex, and translocate into the nucleus where they induce the expression of ISGs (Veals et al. 1992). Type I IFNs induce an antiviral state in stimulated cells and help activate the adaptive immune response (Kumar et al. 2011).

The role of the innate immune system in host defense

The role that these innate immune pathways play in host defense has been revealed by several types of experiments, including (i) forward genetic studies in humans in which the genomes of diseased patients are compared to those of healthy individuals, and a causative gene (in the case of Mendelian diseases that cause primary immunodeficiencies (PIDs)) or associated disease gene(s) (in the case of complex diseases) is identified; (ii) reverse genetic studies in animal models (e.g. *Drosophila melanogaster* or *Mus musculus*) in which a component gene is knocked out, the animal is infected, and various phenotypes of the pathogen (e.g. replication) and/or host (e.g. survival or immune correlates) are quantitated; and (iii) cellular studies in which a component is depleted, the cells are infected, and phenotypes of the pathogen and/or cells are quantitated.

Human genetic studies of PIDs and of common infectious diseases have revealed the role of innate immune components in humans *in vivo*. As explained below, the presentation in knockout mice sometimes phenocopies the human disease, but the human disease often presents with a much more narrow spectrum of infection (Casanova et al. 2011). This may be due to lack of exposure to particular microorganisms that have not revealed particular infectious susceptibilities; environmental differences in humans (e.g. they are being studied *in natura*, i.e. with widespread use of hygienic practices, availability of antibiotics, and vaccination) (Casanova and Abel 2004; Quintana-Murci et al. 2007); the often unnaturally high doses of pathogens given to the mice; or actual physiological differences between these organisms. Furthermore, the lack of clinical disease in humans does not mean that the pathways do not play a role in responding to particular pathogens, but may suggest redundancy in the innate immune system that is compensated for by other pathways *in vivo*.

(i) TLR4 pathway. While no mutations in TLR4 have been identified as a cause of a PID, several mutations have been identified in pathway components downstream of the receptor (Casanova et al. 2011). However, these mutations not only affect the TLR4 pathway but also affect one or many other pathways. Patients with mutations in *MYD88* – affecting most TLR responses (except the TRIF-dependent branch of TLR4 and the TLR3 response) as well as IL-1 receptor responses – suffer from recurrent infection by *Streptococcus pneumonia* leading to meningitis or septicemia, as well as recurrent infection by *Staphylococcus* and *Pseudomonas* (von Bernuth et al. 2008). Patients with mutations in *IRAK4* – affecting TLR and T-cell receptor pathways – suffer from disease that phenocopies *MYD88* deficiency (Picard et al. 2003). Patients with hypomorphic mutations in *NEMO* or hypermorphic mutations in *IκBα* – affecting many NFκB-mediated pathways – suffer from anhidrotic ectodermal dysplasia with immunodeficiency (EDA-ID) leading to infection by pyogenic bacteria (but also to infection by various mycobacteria, viruses, and fungi which are likely unrelated to TLR responses given the *MYD88* and *IRAK4* phenotypes) (Sun et al. 2006). Thus, while the relevance of certain downstream components in protecting against infectious disease in humans is clear, the relevance of the TLR4 pathway in humans is not clear.

In mice, *TLR4* deficiency has been shown to cause increased susceptibility to infection by Gram-negative bacteria, including *Salmonella typhimurium* (Vazquez-Torres et al. 2004; Weiss et al. 2004), and increased susceptibility to endotoxic shock (Poltorak et al. 1998). *Myd88* deficiency has been shown to cause increased susceptibility to a wide range of Gram-negative bacteria, Gram-positive bacteria, viruses, fungi, and parasites, including but not limited to *Streptococcus pneumoniae*, *Staphylococcus aureus*, and

Pseudomonas aeruginosa (von Bernuth et al. 2008). *Irak4* deficiency has been shown to increase susceptibility to infection by *Staphylococcus aureus* (Suzuki et al. 2002). *Nemo* deficiency results in embryonic lethality due to hepatocyte apoptosis (Rudolph et al. 2000) [Makris et al. 2000; Schmidt-Supprian et al. 2000].

(ii) TLR3 pathway. Unlike the TLR4 pathway, the TLR3 pathway is clearly associated with increased susceptibility to infectious disease in humans, though in a surprisingly limited way: i.e. mutations in *TLR3* increase susceptibility to herpes simplex encephalitis (HSE) (Zhang et al. 2007; Guo et al. 2011). Even mutations in *UNC93B1* or in downstream components – *TRIF*, *TRAF3*, *NEMO*, and *TBK1* – which affect multiple pathways besides TLR3, have only been found to increase susceptibility to HSE (Casrouge et al. 2006; Perez de Diego et al. 2010; Audry et al. 2011; Sancho-Shimizu et al. 2011; Herman et al. 2012). *UNC93B1* mutations abolish signaling of TLR3, TLR7, TLR8, and TLR9 (Casrouge et al. 2006; Tabeta et al. 2006). *TRIF* mutations abolish both TLR3 signaling as well as the IFN branch of the TLR4 pathway (Sancho-Shimizu et al. 2011). *TBK1* mutations affect many IFN-inducing pathways (Herman et al. 2012).

In mice, *Tlr3* deficiency has been shown to cause increased susceptibility to infection by several types of RNA viruses including mouse cytomegalovirus (MCMV) (Tabeta et al. 2004), but paradoxically increases resistance to West Nile virus (Wang et al. 2004), influenza A virus (Le Goffic et al. 2006), and phlebovirus (Gowen et al. 2006). *Unc93b1* deficiency has been shown to cause increased susceptibility to MCMV, *Trypanosoma cruzi*, and *Toxoplasma gondii* (Tabeta et al. 2006; Melo et al. 2010; Caetano et al. 2011). *Trif* deficiency has been shown to cause increased susceptibility to MCMV and vaccinia virus (VACV) (Hoebe et al. 2003). *Traf3* deficiency results in early

postnatal lethality (Xu et al. 1996), and *Tbk1* knockout mice die perinatally (Bonnard et al. 2000).

(iii) Cytosolic nucleic-acid sensing pathways. No mutations specifically in the RIG-I or ISD pathways have yet been identified that cause Mendelian infectious disease in humans.

In mice, *RIG-I* deficiency has been shown to cause increased susceptibility to infection by Japanese encephalitis virus (JEV; a positive-sense ssRNA virus) and vesicular stomatitis virus (VSV; a negative-sense ssRNA virus) (Kato et al. 2006), and *RIG-I*-deficient cells are additionally unresponsive to infection by many other ssRNA viruses including influenza A and hepatitis C virus (Kato et al. 2006; Loo et al. 2008). *Mavs* deficiency, which affects both the RIG-I and MDA5 pathways, has been shown to cause increased susceptibility to infection by VSV and EMCV; and *Mavs*-deficient cells are additionally unresponsive to infection by Sendai virus (SV) and Newcastle disease virus (NDV) (Kumar et al. 2006; Sun et al. 2006). *Sting* deficiency has been shown to cause increased susceptibility to infection by herpes simplex virus 1 (HSV-1) and VSV; and *Sting*-deficient cells are additionally unresponsive to infection by SV, CMV, VACV, baculovirus, and *Listeria monocytogenes* (Ishikawa and Barber 2008; Ishikawa et al. 2009).

(iv) Type I (and III) IFN receptor pathway. No mutations in the *IFNAR1* or *IFNAR2* receptors have yet been identified that cause PIDs in humans. However, components downstream of the receptor have been implicated in PIDs. Mutations in *STAT1* can affect type I, type II, and type III IFN signaling as well as signaling of other cytokines and growth factors including IL-6, IL-27, epidermal growth factor (EGF), and platelet-derived growth factor (PDGF) (Casanova et al. 2012). Patients with mutations in *STAT1*

present with Mendelian susceptibility to mycobacterial diseases (MSMD) with increased susceptibility to salmonellosis and several other bacteria, fungi, and parasites (Dupuis et al. 2001), and additionally to viral infection including HSE depending on the pathways affected by the specific mutations in *STAT1* (Dupuis et al. 2003). Mutations in *TYK2* can affect type I and type III IFN signaling as well as signaling of other cytokine including IL-6, IL-10, IL-12, and IL-23 (Casanova et al. 2012). A patient with homozygous mutations in *TYK2* was found to present with atopy, elevated serum IgE levels, and recurrent infection with bacteria, fungi, and viruses including herpes simplex virus (Casanova et al. 2012); while another patient was found to present with the infectious phenotypes, including herpes simplex viral and varicella zoster viral infection, but without atopy or elevated IgE (Kilic et al. 2012). It is not yet clear which of the phenotypes of these patients is due to impaired type I IFN signaling, though the increased infections by herpes simplex virus and varicella zoster virus seem likely given the roles *in vitro* and in animal models (Casanova et al. 2012).

Genetic variants in IFN-stimulated genes have also been associated with increased susceptibility to infection in humans. For example, mutations in the ISG, *ISG15*, cause MSMD (Bogunovic et al. 2012); and rare variants in the ISG, *IFITM3*, have been associated with increased susceptibility to influenza infection (Everitt et al. 2012). Finally, through genome-wide association studies (GWAS) common genetic variants in the type III interferon *IL28B*, which can be induced by viral stimulation of RLR and TLR pathways (Coccia et al. 2004), have been associated with spontaneous clearance of HCV as well as response to treatment with pegylated interferon-alpha (PEG-IFN-a) and

ribavirin (RBV) (Ge et al. 2009; Suppiah et al. 2009; Tanaka et al. 2009; Thomas et al. 2009).

In mice, *Ifnar1* deficiency has been shown to cause increased susceptibility to viruses including VSV, Semliki Forest virus (SFV), VACV, and lymphocytic choriomeningitis virus (LCMV) (Muller et al. 1994). *Stat1* deficiency has been shown to cause increased susceptibility to VSV and *Listeria monocytogenes (Meraz et al. 1996)*. *Tyk2* deficiency has been shown to cause reduced resistance to VACV and MCMV (Karaghiosoff et al. 2000; Strobl et al. 2005). *Jak1* knockout mice die perinatally (Rodig et al. 1998).

These studies establish that variation in innate immune genes can alter the propensity to infectious diseases in humans. However, further studies are needed to better delineate the role of each innate immune pathway in host defense, especially in regards to innate immune genes that have pleiotropic functions in other pathways.

Dysregulation of the innate immune system can lead to autoimmune/autoinflammatory disease

Hyperactive innate immune responses or inappropriate responses against tissues and substances normally considered 'self' could lead to the development of autoimmune disease. While Janeway initially proposed that PAMPs are "characteristic patterns common on infectious agents but absent from the host" (Janeway 1989), it subsequently turned out that the model of PAMPs as foreign is sometimes a gray area, which is demonstrated by several examples below:

1) TLR9 distinguishes bacterial from self DNA by recognizing unmethylated CpG motifs (Krieg et al. 1995; Hemmi et al. 2000), which are present in mammalian genomes but

found in much less frequency than in bacterial genomes (Krieg 2002). Barton *et al*. had shown that localization of TLR9 to endolysosomes is also important for distinguishing foreign from self DNA (Barton et al. 2006). When they altered TLR9 localization to the cell surface, the re-localized TLR9 could recognize self DNA. Barton *et al*. argued that TLR9 localization inside the cell, as well as the expression of DNases outside of the cell that degrade extracellular DNA, prevent self DNA recognition.

2) Stetson *et al*. had shown that transfection of dsDNA of any random sequence into mammalian cells can stimulate the ISD pathway (Stetson and Medzhitov 2006). It is not yet entirely clear how the ISD pathway distinguishes foreign from self DNA, but it is believed that the mechanism might be receptor localization to the cytosol, where DNA is not normally found. This is not a perfect explanation because in certain normal cellular states – such as during cell division when the nuclear envelope breaks down – genomic DNA is exposed to the cytosol; however, the genomic DNA is coated with proteins such as nucleosomes which could possibly affect detection. Also others have argued that viral DNA recognition may occur in the nucleus (Kerur et al. 2011; Li et al. 2012; Orzalli et al. 2012), though they have not explained how the genomic DNA evades recognition.

3) As mentioned above, host-derived signals of cellular stress (e.g. uric acid and ATP), which are referred to as danger-associated molecular patterns (DAMPs), can be recognized by the NLR family of the innate immune system.

Genetic variants that cause breakdown of the innate immune system's ability to distinguish foreign vs. self, that decrease the function of negative regulators of the innate immune response, or that increase the function of positive regulators of the innate immune

response have been found to alter the propensity to certain autoimmune diseases, as described below.

(i) TLR pathways. GWA studies have identified variants near several innate immune genes that associate with systemic lupus erythematosus (SLE) in humans, including *IRAK1*, *TNFAIP3*, and *IRF5* (Graham et al. 2007; Graham et al. 2008; Jacob et al. 2009; Stahl et al. 2010). GWAS has also identified variants near *TNFAIP3*, *IRF5*, *NFKBIE*, and *REL* that associate with rheumatoid arthritis (Plenge et al. 2007; Gregersen et al. 2009; Stahl et al. 2010; Okada et al. 2012). TNFAIP3 is a negative regulator of TLR responses that controls ubiquitination of key signaling proteins including TRAF6 (Boone et al. 2004; Wertz et al. 2004); in mice, *Tnfaip3* deficiency leads to the spontaneous development of multiorgan inflammation and severe cachexia (Lee et al. 2000; Boone et al. 2004; Turer et al. 2008). IRF5 is an IRF family transcription factor; in mice, *Irf5* deficiency impairs pro-inflammatory cytokine production downstream of *Myd88* (Takaoka et al. 2005). *IRAK1* is an interleukin-1 receptor associated kinase involved in Toll/IL-1 receptor signaling; *REL* is an NFκB family member; and *NFKBIE* (IκB-ϵ) is an inhibitor of NFκB as described above.

Again, these genes are not specific to the TLR pathway and affect other pathways in both the innate and adaptive immune system; thus, it is difficult to determine the contribution of altered TLR responses to these autoimmune diseases. However, there is substantial evidence from mouse models that TLR responses (especially TLR7 and TLR9 responses) play a role in the pathogenesis of autoimmune diseases such as SLE and psoriasis (Leadbetter et al. 2002; Pisitkun et al. 2006; Lande et al. 2007; Marshak-

Rothstein and Rifkin 2007), suggesting that altered TLR responses caused by the human variants may explain at least part of the clinical associations.

GWA studies have also identified variants near the TLR-stimulated immunoregulatory cytokine, *IL10*, that associate with type I diabetes mellitus (T1DM) (Barrett et al. 2009). In addition to the GWA studies, Mendelian mutations in *IL10* cause infant colitis in humans (Glocker et al. 2010); in mice, *Il10* deficiency results in a similar phenotype, which is rescued by knockout of *Myd88*, demonstrating that Toll/IL-1 signaling is essential to the inflammatory phenotype in the animal model (Kuhn et al. 1993; Rakoff-Nahoum et al. 2006).

(ii) Cytosolic nucleic-acid sensing pathways. Variants in the RLR, *MDA5*, have been associated with protection from type I diabetes mellitus (T1DM) (Smyth et al. 2006; Nejentsev et al. 2009). Some of the *MDA5* variants are believed to be deleterious to the function of the protein (Shigemoto et al. 2009), suggesting that the innate immune response to viral infection (e.g. by enteroviruses, which had previously been associated with T1DM (Hyoty and Taylor 2002)) may increase propensity to T1DM.

Hyperactivation of the ISD pathway is believed to cause certain forms of autoimmune disease, demonstrated by mutations in a regulatory gene called *TREX1*. TREX1 is an endoplasmic reticulum (ER)-localized 3'->5' exonuclease. This gene was initially discovered by linkage studies as a cause of Aicardi-Goutières syndrome (AGS), a family of rare monogenic disorders characterized by spontaneous (non-infectious) encephalitis in humans and myocarditis in mice (Crow et al. 2006). Yang *et al.* showed that AGS patients with mutations in TREX1 accumulate ssDNA in the cytoplasm (Yang et al. 2007). Stetson *et al.* hypothesized that the cytoplasmic DNA may lead to aberrant

activation of the ISD pathway (Stetson et al. 2008). They found that knockout of IFNAR1 or of the ISD pathway components STING or IRF3 rescued the disease phenotype in *Trex1*-/- mice (Stetson et al. 2008; Gall et al. 2012). These results confirmed that *Trex1* is in epistasis with ISD pathway components. Stetson *et al.* sequenced the ssDNA that accumulates in the cytoplasm, found that the DNA is enriched for retroelement DNA, and proposed a model that explains the pathogenesis: retroelement DNA can normally exit the nucleus during its life cycle, *Trex1* normally degrades the DNA before recognition by the ISD receptor, and in the absence of TREX1 the DNA can accumulate in the cytoplasm and lead to constitutive activation of the ISD pathway (Stetson et al. 2008). In line with this hypothesis, Beck-Engeser *et al*. showed that treatment of *Trex1*-/- mice with reverse transcriptase inhibitors could rescue the disease phenotype (Beck-Engeser et al. 2011).

Because there is a phenotypic overlap of AGS with systemic lupus erythematosus (SLE), Lee-Kirsch *et al.* sequenced the *TREX1* gene in patients with SLE and found that variants in the *TREX1* gene – some of which are the same as the AGS mutations except occur monoallelically – are associated with SLE (Lee-Kirsch et al. 2007a). These variants are found in ~2% of the SLE patients that they sequenced, suggesting that overactivation of the ISD pathway may occur in at least 2% of SLE cases.

TREX1, also known as *DNase III*, is one of the four DNases in humans, which include: *DNASE1*, *DNASE2*, *TREX1* (*DNase III*), and *FEN1* (*DNase IV*). Mutations in other DNase genes are also associated with autoimmune disease. Mutations in DNASE1 were found in patients with SLE (Yasutomo et al. 2001); in mice, knockout of *Dnase1* causes an SLE-like phenotype, including glomerulonephritis and the presence of anti-

nuclear antibodies (Napirei et al. 2000). Knockout of *Dnase2* and *Fen1* also cause autoimmune disorders in mice (Kawane et al. 2001; Yoshida et al. 2005; Kawane et al. 2006; Zheng et al. 2007).

(iii) Type I IFN pathway. Genetic variants in ISGs have been found to be associated with autoimmune diseases. For example, in addition to *TREX1* mutations, Mendelian mutations in the ISG, *ADAR1*, can also cause AGS (Rice et al. 2012). Variants near the ISG, *STAT4*, have been associated with SLE and RA in GWA studies (Remmers et al., 2007). Additionally, signatures of IFN upregulation have been observed in several types of autoimmune disorders such as SLE, AGS, and spondyloenchondrodysplasia, suggesting a pathogenic role (Bennett et al. 2003; Crow 2011).

(iv) NLR pathways. Although the NLR pathways are not highlighted here, various Mendelian mutations in NLR components can cause autoinflammatory syndromes (Masters et al. 2009) including: NOD2 mutations that cause Blau syndrome; NLRP3 mutations that cause Muckle-Wells syndrome, Familial cold urticaria, or Neonatal onset multisystem inflammatory disease (NOMID); MEFV (pyrin) mutations that cause Familial Mediterranean Fever (FMF); PSTPIP1 mutations that cause Pyogenic sterile arthritis, pyoderma gangrenosum, acne (PAPA); and IL1RN mutations that cause Deficiency of the interleukin-1–receptor antagonist (DIRA). Additionally, variants in *NOD2* are associated with Crohn's disease (Cho 2008), and the NLRP3 pathway is believed to recognize the pathologic uric acid crystals in patients with gout and pseudogout (Martinon et al. 2006). IL-1β antagonists such as anakinra (a soluble IL-1 receptor) have shown promise in many of these diseases (Schroder and Tschopp 2010).

Taken together, these studies establish that variation in innate immune genes can alter propensity to autoimmune/autoinflammatory diseases in humans. Like the infectious diseases, in certain cases further studies are needed to understand the contribution of specific innate immune pathways.

Characterization of variation in the human innate immune system

The clinical phenotypes described above clearly demonstrate that innate immune responses vary between individuals. The Mendelian mutations demonstrate that certain rare variants in innate immune components cause infectious diseases or autoimmune/autoinflammatory diseases. The GWAS and candidate association studies demonstrate that certain common (minor allele frequency > 1%) variants in innate immune components associate with infectious or autoimmune/autoinflammatory diseases. **But because the clinical phenotypes and genetic studies have not been comprehensive by any means, there is likely much more variation in the human innate immune system that affects propensity to infectious and autoimmune diseases.**

Thus, in addition to forward genetic studies, researchers have more directly characterized variation in the innate immune system by: (1) sequencing innate immune genes in large numbers of individuals; and (2) stimulating cells from large numbers of individuals with innate immune ligands to compare their responses. Moreover, researchers have characterized genome-wide expression variation in immune cells – such as monocytes, monocyte-derived dendritic cells (MoDCs) and B cells – and associated this phenotypic variation with genome-wide genotype data (Barreiro et al. 2012; Fairfax et al. 2012). This has resulted in the identification of eQTLs (expression quantitative trait loci), which are genetic variants that associate with variation in the

expression of gene(s), some of which influence the expression of innate immune components though this has not been explicitly examined.

(i) TLR pathways. Barreiro *et al*. sequenced the TLR genes in 158 healthy individuals of various populations (Barreiro et al. 2009). They demonstrated that the nucleic acid-sensing TLRs (TLR3, TLR7, TLR8, TLR9) have evolved under strong purifying selection, suggesting that they serve essential, nonredundant functions. In contrast, they found that the cell surface TLRs display relatively high rates of nonsynonymous variations in the population, suggesting higher immunological redundancy. Fornarino *et al*. sequenced the TIR domain-containing adaptor genes (including *MYD88* and *TRIF*) in 183 healthy individuals of various populations (Fornarino et al. 2011). They demonstrated that *MYD88* and *TRIF* have also evolved under strong purifying selection (suggesting that they serve essential, nonredundant functions) and have been subject to positive selection suggesting adaptation during human history.

Wurfel *et al*. stimulated whole blood from 102 healthy individuals with LPS, and then measured production of a panel of cytokines by ELISA (Wurfel et al. 2005). They observed substantial variation in cytokine production, though could not distinguish genetic from technical or biological confounders. Hedl *et al*. stimulated the NOD2 and TLR2 receptors in monocyte-derived cells from 77-102 healthy individuals, and measured production of cytokines by ELISA (Hedl and Abraham 2012). They also observed substantial variation in cytokine production, and found that genetic variants in *IRF5* associate with part of this phenotypic variation.

(ii) Cytosolic nucleic-acid sensing pathways. Vasseur *et al*. sequenced the RLR genes in 186 healthy individuals of various populations (Vasseur et al. 2011). They

demonstrated that *RIG-I* is the most genetically constrained RLR (i.e. low tolerance of amino acid-altering variation, and low levels of nucleotide diversity and population differentiation), suggesting an essential role in host survival. They also found that certain variants in *MDA5* and *LGP2* have been subject to positive selection, suggesting a selective advantage of these mutations to certain populations.

Jin *et al*. genotyped the *STING* gene in 1074 individuals. Their data revealed a haplotype in *STING* that contains three non-synonymous single nucleotide polymorphisms (SNPs), R71H-G230A-R293Q, that cause a >90% reduction of STING function (Jin et al. 2011). ~15% of Asians and ~1.7% of Europeans are homozygous for this haplotype.

(iii) Type I (and type II and III) IFN pathways. Manry *et al*. sequenced the types I, II, and III IFN genes in 186 healthy individuals of various populations (Manry et al. 2011). They demonstrated that IFNγ (the type II IFN) and certain IFNα subtypes (notably *IFNA6*, *IFNA8*, *IFNA13*, and *IFNA14*) have evolved under strong purifying selection, suggesting that they serve essential, nonredundant functions. They found that other type I IFNs, including *IFNA10* and *IFNE*, have high frequencies of non-synonymous mutations, suggesting redundancy of these genes. Finally, they found that type III IFNs are the only IFNs to show evidence of positive selection, suggesting a selective advantage of certain genetic variants to certain populations.

These studies have characterized additional variation in innate immune genes in humans, some of which may affect clinical phenotypes. For example, some of the variants in the type III IFN genes that have been subject to positive selection overlap with variants found in GWA studies of viral disease (Manry et al. 2011). In addition, variation in the NOD2 and TLR2

responses were linked to genetic variants found in GWA studies of autoimmune disease (Hedl and Abraham 2012).

Overview of objectives

Aim #1: Identify novel ISD pathway components. Our first objective was to identify novel components of the ISD pathway, since it is one of the most poorly understood pathways in the innate immune system and is clearly linked to human disease. Because homology-based and candidate gene approaches had not been yielding further discoveries, we used an unbiased approach that focused on candidates from proteomic and genomic experiments. We then developed a high-throughput siRNA screen to functionally test our candidates for a potential role in the ISD pathway.

Aim #2: Characterize variation in innate immune responses in primary dendritic cells in humans, and determine whether the functional variation is caused by genetic variation. Our second objective was to characterize variation in innate immune responses to LPS, influenza, and IFNβ in monocyte-derived dendritic cells of a large number of healthy individuals. These ligands would allow us to characterize variation in the TLR4, RIG-I and TLR3, and IFNAR pathways, respectively. We sought to determine whether the functional variation is caused by genetic variation by using serially-replicated samples to control for technical and biological noise, and by associating genotype data with our expression and response phenotypes.

CHAPTER 1: AN INTEGRATIVE APPROACH IDENTIFIES CRITICAL REGULATORS OF THE INNATE IMMUNE RESPONSE TO CYTOSOLIC DNA AND RETROVIRAL INFECTION

Mark N. Lee, Matthew Roy, Shao-En Ong, Philipp Mertins, Alexandra-Chloé Villani, Weibo Li, Farokh Dotiwala, Jayita Sen, John G. Doench, Megan H. Orzalli, Igor Kramnik, David M. Knipe, Judy Lieberman, Steven A. Carr, Nir Hacohen

AUTHOR CONTRIBUTIONS

Lee MN is responsible for performing the DNA and protein pulldowns, helping to optimize the RNAi screen conditions, helping to perform the RNAi screen, performing experiments with small molecules, performing cDNA rescue experiments, performing *Abcf1* experiments, performing experiments with retrovirus, performing experiments with patient cells, and writing the manuscript.

Roy M is responsible for helping to optimize the RNAi screen conditions, helping to perform the RNAi screen, performing experiments in knockout cells, and performing experiments with HSV-1 *d109*.

Ong SE is responsible for performing the mass spectrometry of the DNA SILAC experiment and analyzing this data.

Mertins P is responsible for performing the mass spectrometry of the ABCF1-HA and STING-HA SILAC experiments and analyzing these data.

Villani AC is responsible for analyzing the microarray data and helping to revise the manuscript.

Li W is responsible for culturing cells and helping to clone DNA constructs.

Dotiwala F is responsible for performing the immunofluorescence and co-immunoprecipitation experiments.

Sen J is responsible for helping to perform the HSV-1 *d109* infections.

Doench JG is responsible for providing the pCW57d-P2AR vector.

Orzalli MH is responsible for providing the HSV-1 *d109* virus.

Kramnik I is responsible for providing the *Sp110* mutant cells.

Knipe DM is responsible for supervising the HSV-1 *d109* experiments.

Lieberman J is responsible for supervising the immunofluorescence and co-immunoprecipitation experiments.

Carr SA supervised the SILAC mass spectrometry experiments.

Hacohen N supervised the entire project and manuscript writing.

ACKNOWLEDGMENTS

We are grateful to Y.J. Crow for AGS patient fibroblasts and control cells; T. Lindahl for *Trex1*-/- MEFs and wild type control MEFs; D. Sabatini and D. Kwiatkowski for *p53*-/- MEFs; N.A. DeLuca for HSV-1 *d*109 virus; C. Shamu at ICCB (HMS) for the siRNA library and expert advice; the RNAi Consortium (TRC) at the Broad Institute for assistance with siRNA screening; J. Qiao for assistance with mass spectrometry; M. Rooney and C. Ye for advice about statistical analyses; J. Kagan, N. Haining, L. Glimcher, and M. Brenner for valuable discussions; and W.F. Pendergraft III and other members of the Hacohen laboratory for critical review of the manuscript. This work was supported by the NIH Director's New Innovator award (N.H.). M.N.L. is supported by an NIH MSTP fellowship; F.D. by a GSK-IDI fellowship; and J.L. by NIAID grant AI102816.

ABSTRACT

The innate immune system senses viral DNA that enters mammalian cells – or in aberrant situations, self DNA – and triggers type I interferon production. Here we present an integrative approach that combines quantitative proteomics, genomics, and small molecule perturbations to identify novel genes involved in this pathway. We silenced 809 candidate genes, measured the response to dsDNA, and connected resulting hits with the known signaling network. We identified ABCF1 as a critical protein that associates with dsDNA and the known DNA-sensing components, HMGB2 and IFI16. We also found that CDC37 regulates stability of the signaling molecule, TBK1, and that chemical inhibition of CDC37 and several other pathway regulators potently modulates the innate immune response to DNA and retroviral infection.

INTRODUCTION

The innate immune system detects viral infection primarily by recognizing viral nucleic acids inside an infected cell (Barbalat et al. 2011). In the case of retroviruses, which are RNA viruses that replicate via a DNA intermediate, genomic RNA can be recognized in the endosomes of specialized innate immune cells using the receptor TLR7 (Altfeld et al. 2011), while the reverse transcribed DNA is believed to be recognized during entry into a host cell by a cytoplasmic DNA sensor(s) that triggers type I IFN production (Yan et al. 2010). This latter response has been observed in cells that lack the endoplasmic reticulum (ER)-associated 3'->5' exonuclease, TREX1. When TREX1 is present, it can degrade the viral DNA before sensing occurs.

A similar fate appears to be the case with self DNA from the host cell, as deficiency in *Trex1* leads to the accumulation of endogenous retroelements and genomic DNA in the cytoplasm, causing aberrant overactivation of the DNA-sensing pathway and subsequent initiation of autoimmune disease (Stetson et al. 2008; Gall et al. 2012). Indeed, deleterious variants in *Trex1* cause the Mendelian autoimmune disorders, Aicardi-Goutières syndrome (AGS) (Lee-Kirsch et al. 2007b) and familial chilblain lupus (Lee-Kirsch et al. 2007a) in humans, and are associated with a subset of systemic lupus erythematosus (SLE) (Lee-Kirsch et al. 2007b). Genetic ablation of DNA-sensing pathway components (i.e. *Irf3* or *Sting*) or the type I IFN receptor can preclude disease onset in animal models (Stetson et al. 2008; Gall et al. 2012), demonstrating the key epistatic relationship between *Trex1* and the interferon stimulatory DNA pathway and linking the recognition of DNA viral and retroviral infection to the initiation of autoimmune disease. Identifying components of this pathway may facilitate the identification of drugs that modulate the DNA-sensing response to mitigate disease.

The mechanism by which cytosolic DNA elicits production of type I IFNs is a long-standing problem (Isaacs et al. 1963; Jensen et al. 1963; Rotem et al. 1963) that has gained key insights more recently. Several of the downstream components of this pathway are believed to be shared with the RIG-I pathway, which mediates an innate immune response to foreign cytosolic RNA (Yoneyama et al. 2004). Specific DNA sequences, e.g. AT-rich DNA, are recognized by cytoplasmic RNA polymerase III, which transcribes the DNA ligand, and then RIG-I recognizes the resulting RNA product (Ablasser et al. 2009; Chiu et al. 2009). However, the innate immune recognition of retroviruses or retroelements does not seem to involve this latter pathway (Yan et al. 2010; Gall et al. 2012). Instead, cytosolic DNA can be recognized in a sequence-independent manner and this latter pathway does not use the receptor RIG-I or the RIG-I adaptor, MAVS (Ishii et al. 2006; Stetson and Medzhitov 2006; Sun et al. 2006). Instead, it relies on the intracellular transmembrane protein, STING (Ishikawa and Barber 2008; Ishikawa et al. 2009), which binds to the kinase TBK1 and the transcription factor IRF3 to allow IRF3 phosphorylation (Tanaka and Chen 2012). Phosphorylated IRF3 then dimerizes and translocates into the nucleus to induce *Ifnb1* expression. In addition, certain HMGB family proteins as well as the AIM2-like receptor, IFI16, are believed to play at least a partial role (Yanai et al. 2009; Unterholzner et al. 2010). However, the mechanism is still not fully understood.

We describe here an integrative approach for identification of novel components of this DNA-sensing pathway (referred to as the 'interferon stimulatory DNA (ISD) pathway' below) and the innate immune response to retroviral infection. We combined unbiased quantitative proteomics with curation of candidates from existing proteomic, genomic, and domain-based datasets, and functionally tested 809 of these candidates by high-throughput loss-of-function screening. We then validated hits by chemical inhibition, cDNA rescue, or targeted knockout,

and mined existing protein-protein interaction datasets (host-host and host-viral) to form a network model of the ISD pathway.

RESULTS

Generation of a candidate gene set by curation and quantitative proteomics

We generated a set of candidate genes from proteomic, genomic, and domain-based datasets that we hypothesized contain unidentified ISD pathway components (**Fig. 1.1a**). First, we used protein-protein interaction (PPI) datasets to nominate 36 candidate proteins that interacted with the known DNA-sensing signaling proteins STING (Ishikawa and Barber 2008; Ishikawa et al. 2009), TBK1 (Ishii et al. 2006; Tanaka and Chen 2012), IKKϵ (Ishii et al. 2006), and IRF3 (Stetson and Medzhitov 2006) from a recent mass spectrometry study (Li et al. 2011), as well as 99 candidates from our own mass spectrometry-based list of putative STING-interacting proteins (**Supplementary Table 1.1**). Second, we selected 321 DNA- and interferon-stimulated genes (ISGs) from our own and existing microarray datasets (see **Materials and Methods**) based on the hypothesis that a subset of components of this pathway are feedback-regulated (Tsuchida et al. 2010; Unterholzner et al. 2010). Third, we focused on 126 annotated phosphatases (Gene Ontology (GO):0004721) and 71 deubiquitinases (GO:0004221 and ref. (Nijman et al. 2005)) as part of our pilot screen to identify regulators of the ISD pathway (Tsuchida et al. 2010; Tanaka and Chen 2012).

Because there was no existing dataset of cytoplasmic DNA-interacting proteins, we used quantitative proteomics to discover such proteins. We prepared cytoplasmic extracts from mouse embryonic fibroblasts (MEFs) (**Supplementary Fig. 1.1a**), and added biotinylated 45 base pair double-stranded DNA ('ISD' sequence (Stetson and Medzhitov 2006)) coupled to streptavidin beads as bait. We utilized three-state SILAC (stable isotope labeling by amino acids in cell culture) to label and quantitate peptides using mass spectrometry (Ong et al. 2002), with medium isotope-labeled cells used for a negative control (beads alone), light isotope-labeled

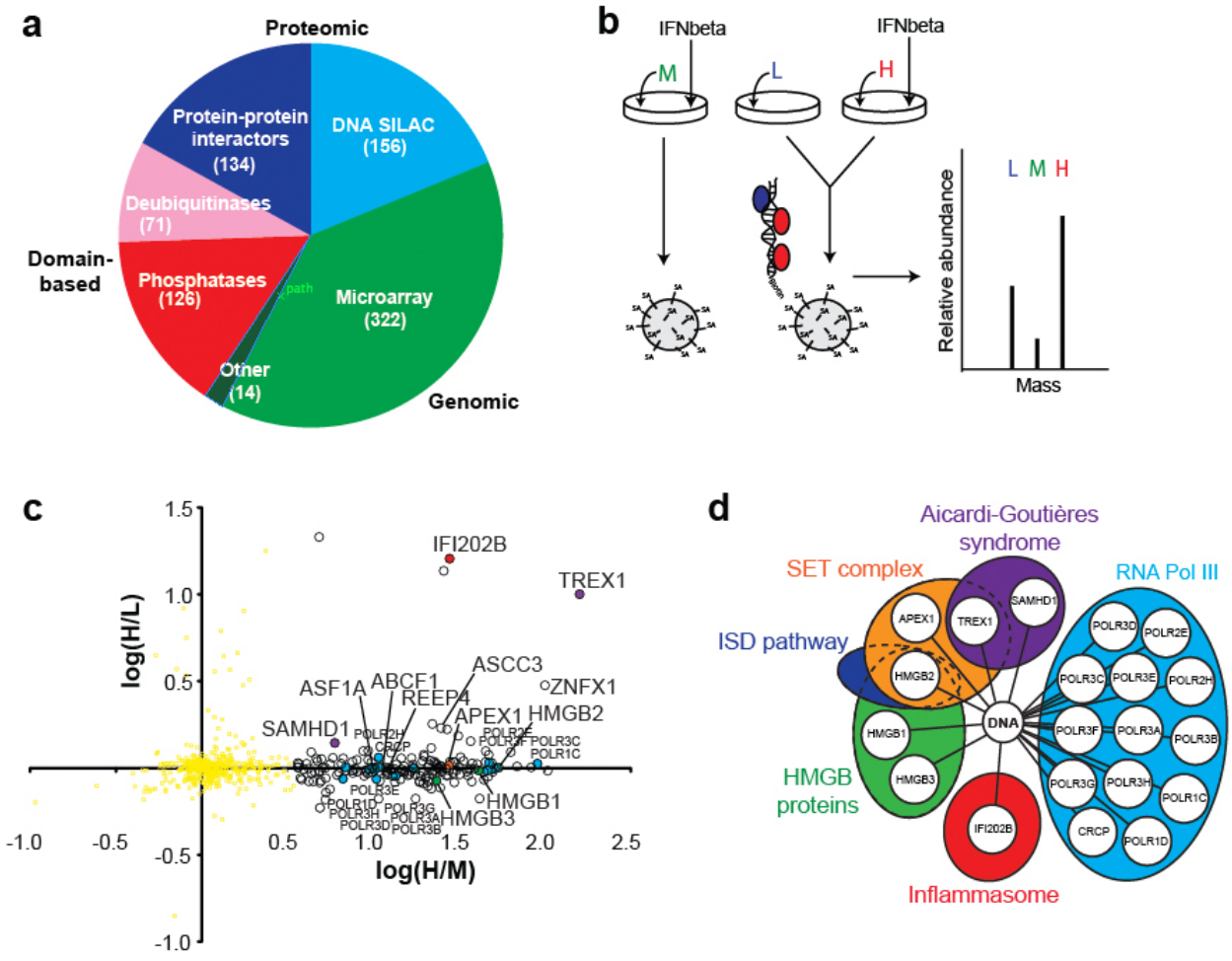


Figure 1.1. Generation of a candidate gene set by curation and quantitative proteomics.
(**a**) Proteomic, genomic, and domain-based gene sets that we hypothesized contain unidentified ISD pathway components were curated or experimentally derived. (**b**) Schematic of DNA-interacting SILAC experiment. MEFs were labeled with light (L)-, medium (M)-, or heavy (H)-isotope SILAC solutions. Cells were pre-treated with IFNβ or left unstimulated. Cytoplasmic extracts were prepared and incubated with or without biotinylated ISD. ISD was precipitated with streptavidin beads, and precipitated proteins were trypsinized and subjected to mass spectrometry. (**c**) DNA-interacting proteins. The x-axis (H/M) corresponds to ISD interaction, and the y-axis (H/L) corresponds to IFNβ regulation. Open circles, ISD-interacting hit; yellow squares, non-hit; selected hits are labeled. Open circles are colored to correspond with pathways in (d). (**d**) DNA-interacting SILAC hits span several known pathways involved in sensing cytosolic DNA.

cells for bead-DNA precipitation, and heavy isotope-labeled cells for bead-DNA precipitation preceded by IFNβ stimulation to upregulate pathway components (**Fig. 1.1b**).

While only a handful of bands were visually distinguishable by protein electrophoresis (**Supplementary Fig. 1.1b,c**), we identified 184 proteins with SILAC ratios that showed enrichment for DNA binding following mass spectrometry (**Fig. 1.1c**, **Supplementary Table 1.2**, and **Supplementary Fig. 1.1d**). Among the 184 proteins, 121 (64.2%) were classified by Gene Ontology as having nucleic acid binding function ($P = 5.95 \times 10^{-58}$; GO:0003676), and others were components of DNA-binding complexes.

In total, twenty of the identified proteins (10.9%) represent the majority of known players involved in the immune sensing of cytosolic DNA (**Fig. 1.1d**). We identified known components of DNA sensing pathways including: the HMGB family proteins (HMGB1, HMGB2, HMGB3) (Yanai et al. 2009), components of the AIM2 inflammasome (IFI202B and the HMGB proteins) (Roberts et al. 2009; Yanai et al. 2009), and the cytosolic RNA polymerase III complex (POLR3A, POLR3B, POLR3C, POLR3D, POLR3E, POLR3F, POLR3G, POLR3H, POLR1C, POLR1D, POLR2E, POLR2H, and CRCP). We identified three members of the SET complex (TREX1, APEX1, and HMGB2) that regulate the ISD pathway as well as HIV-1 detection and infection (Stetson et al. 2008; Yan et al. 2009; Yanai et al. 2009; Yan et al. 2010). We also identified associated proteins responsible for the autoimmune disease, Aicardi-Goutières syndrome (SAMHD1 and TREX1 (Crow et al. 2006; Rice et al. 2009), which are involved in regulating retroviral and retroelement detection (Stetson et al. 2008; Laguette and Benkirane 2012)). Our findings validate the utility of quantitative mass spectrometry as an approach to find known components of cytosolic DNA sensing pathways, and we hypothesized that the dataset may also contain novel DNA-sensing components.

High-throughput loss-of-function screening of candidates and network analysis

In total, we generated a list of 809 proteomic, genomic, or domain-based candidates (that were also available in the Dharmacon SMARTpool siRNA library) to test as potential components of the ISD pathway. We developed a robust high-throughput siRNA (small interfering RNA) screening assay (**Supplementary Fig. 1.2a**) in which we knocked down the genes in MEFs, stimulated the wells with transfected dsDNA (ISD), and measured production of the IFN-inducible protein, CXCL10, by ELISA (**Fig. 1.2a**). We used CXCL10 here as the representative of the set of IFN-inducible genes because of its high level of induction in response to nucleic acids, its dependence on IRF3, and the availability of robust protein and RNA assays (Okabe et al. 2009). We measured cell survival after knockdown to control for potential cytotoxic effects of the siRNAs. The averages of triplicate wells are shown in **Figure 1.2b** and **Supplementary Tables 1.3a-e**. Knockdown of selected hits was confirmed by qRT-PCR (**Supplementary Fig. 1.2b**).

We hypothesized positions for selected hits in the ISD pathway by bringing together information from our own and existing PPI datasets (**Fig. 1.2c** and **Supplementary Methods**) (Bouwmeester et al. 2004; Li et al. 2011). We then validated the knockdown phenotypes of several of the screening hits using targeted knockouts, cDNA rescue, and chemical inhibition, as described below.

Validation by targeted knockout, cDNA rescue, and chemical inhibition

The ISD signaling pathway can be divided broadly into three main processes: DNA sensing, primary signaling, and secondary (IFN) signaling (**Fig. 1.2c**). At the level of DNA sensing, several of our cytoplasmic DNA-interacting mass spectrometry hits were found to have

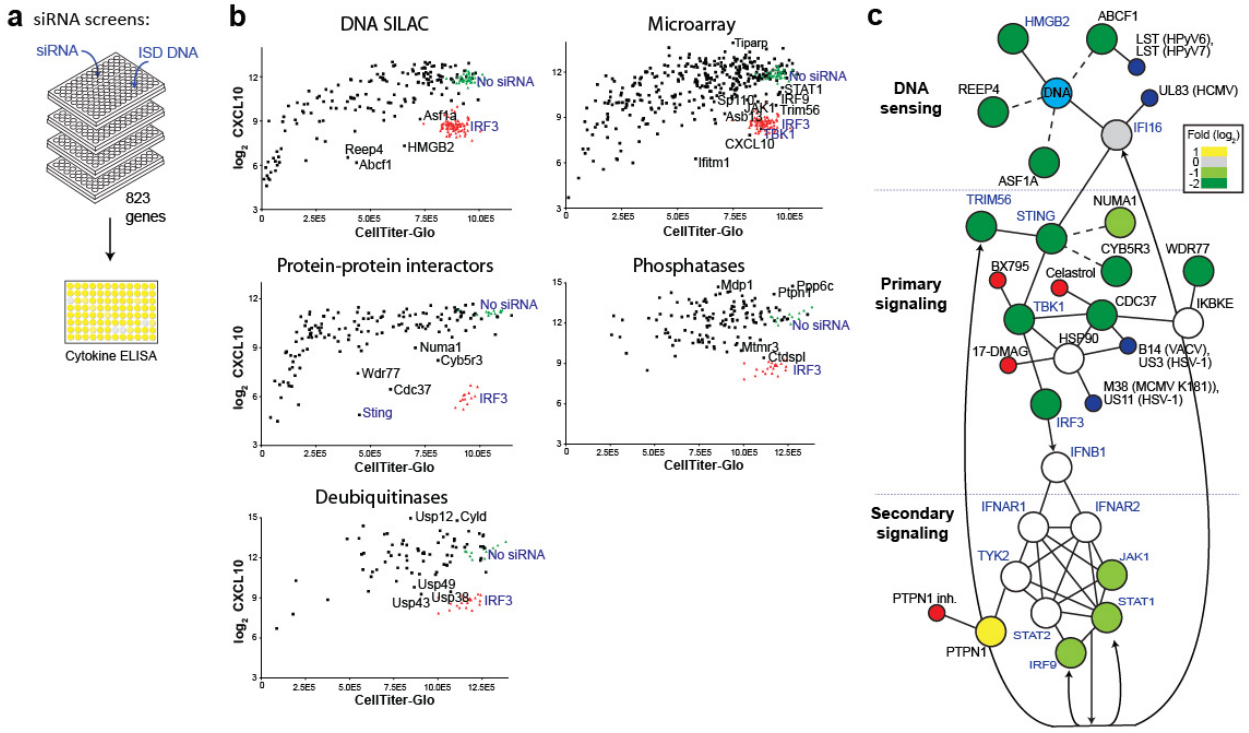


Figure 1.2. High-throughput loss-of-function screening and network analysis of hits.
(**a**) Schematic of siRNA screening assay. Following siRNA knockdown, each well was stimulated with ISD, and CXCL10 was measured in the supernatants by ELISA. (**b**) siRNA screen results. Log_2 CXCL10 (pg/mL) is graphed on the y-axis. Cell viability after knockdown was measured by CellTiter-Glo; relative luminescence units are graphed on the x-axis. The averages of three replicate wells for each gene are represented as black circles; red triangles at the bottom-right of each graph represent siIRF3 positive controls; green triangles at the top-right of each graph represent wells with no siRNA. Selected genes are labeled; control genes are labeled with blue text. (**c**) Interaction among known and novel proteins involved in the ISD network, assessed using PPI datasets. Blue text, known ISD pathway components. Solid lines, interactions defined in other datasets; dashed lines, interactions defined in our datasets; arrows, transcriptional regulation. Green/gray/yellow circles, screened by siRNA and colored by log_2(fold) values relative to no siRNA control; white circles, toxic siRNA or not screened; blue circles, viral protein; red circles, chemical inhibitor.

functional phenotypes in our siRNA screen, including *Hmgb2* and *Abcf1* (**Supplementary Table 1.3a**). HMGB2 is a known nucleic acid sensing pathway member that interacts with DNA (Yanai et al. 2009). Knockout of *Hmgb2* reduced ISD-induced *Ifnb1* and *Cxcl10* production 2-3 fold (**Supplementary Fig. 1.3a**), consistent with the role of *Hmgb2* in the response to B-DNA (poly(dA-dT)-poly(dT-dA)) which stimulates the ISD and RNA polymerase III pathways (Ablasser et al. 2009; Chiu et al. 2009; Yanai et al. 2009). We validated the *Abcf1* phenotype using additional siRNAs and cDNA rescue. Using 14 different siRNAs targeting *Abcf1*, we measured *Abcf1* mRNA expression as well as CXCL10 induction in response to ISD stimulation. Knockdown of *Abcf1* correlated with CXCL10 induction ($R^2 = 0.62$), with the screening siRNA pool (si-0) and two other siRNAs (si-1 and -2) inhibiting both *Abcf1* mRNA and protein expression and CXCL10 induction most strongly (**Fig. 1.3a**). We validated the strongest *Abcf1* siRNA (si-1) by cDNA rescue. We first created an siRNA-resistant cDNA (*Abcf1*(rescue) gene; **Fig. 1.3b**). We cloned this cDNA into a tet-on lentiviral vector (**Supplementary Fig. 1.3b**), transduced the construct into MEFs, and repeated siRNA-mediated knockdown in the presence of varying amounts of doxycycline to titrate the expression of the cDNA. We stimulated the cells with ISD and read out CXCL10 production by ELISA. Knockdown of *Abcf1* reduced CXCL10 production by 14.9-fold ($P < 0.01$; **Fig. 1.3c**). Expression of *Abcf1*(rescue) cDNA, but not of a *Renilla* cDNA control, significantly rescued this phenotype in a dose-dependent manner ($P < 0.001$; **Fig. 1.3c** and **Supplementary Fig. 1.3c**). *Abcf1* is a cytosolic and ER-localized member of the ATP-Binding Cassette (ABC) family of transporters with a role in translational control (Paytubi et al. 2009), but unlike other members of the ABC family, the Abcf subfamily genes lack transmembrane domains. Although a role in DNA sensing had not been previously observed, there is evidence that human polyomavirus 6 and 7 proteins interact with ABCF1

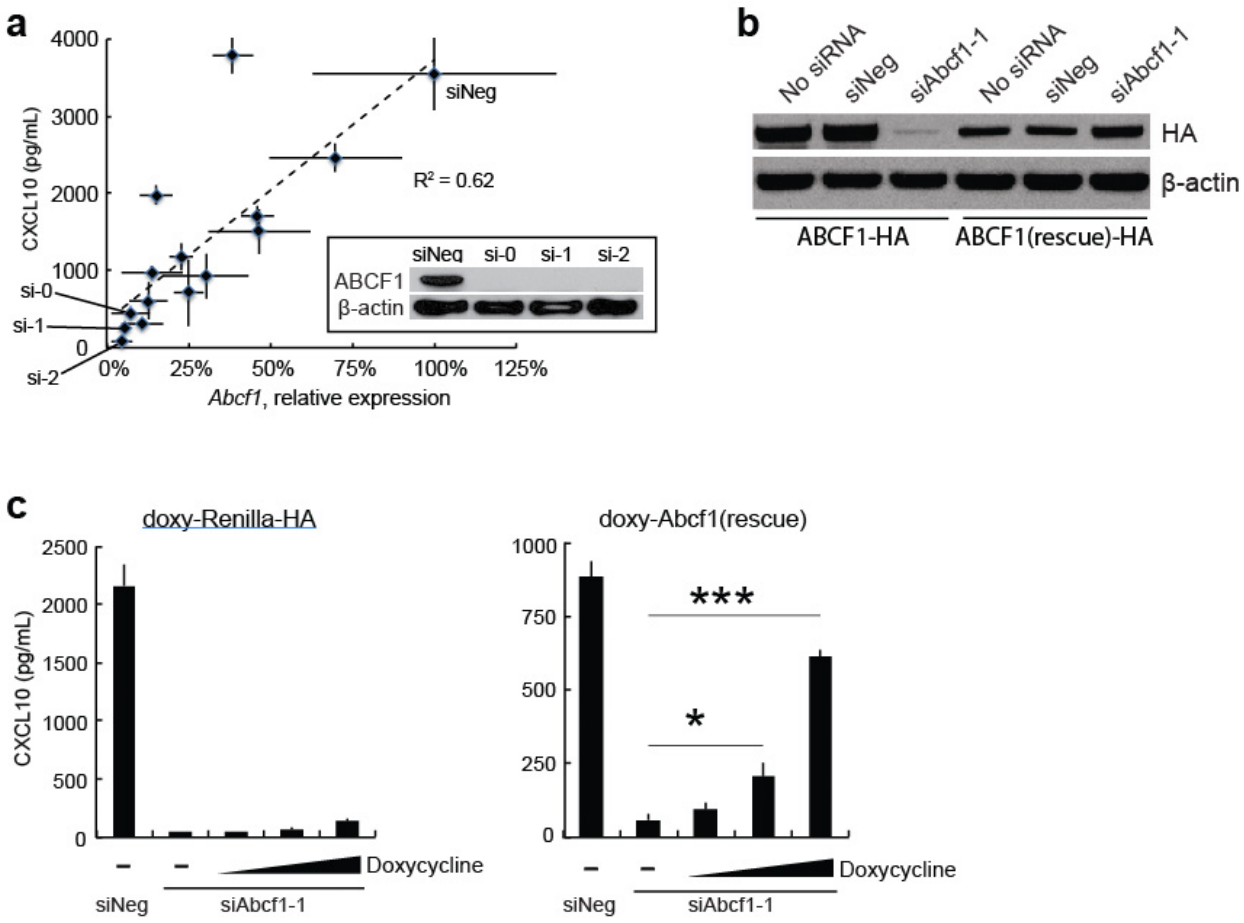


Figure 1.3. Validation of the screening hit, *Abcf1*. (**a**) Knockdown-phenotype correlation. MEFs treated with control siRNA (siNeg) or 14 different siRNAs targeting *Abcf1* were stimulated with ISD. *Abcf1* mRNA expression was measured by qRT-PCR, and CXCL10 production was measured by ELISA. Inset: MEFs were treated with siNeg, siAbcf1-0, siAbcf1-1, or siAbcf1-2, and lysates were immunoblotted with anti-ABCF1 and anti-β-actin antibodies. (**b**) MEFs stably expressing *Abcf1*-HA or siAbcf1(si-1)-knockdown-resistant *Abcf1*(rescue)-HA were treated with siAbcf1(si-1), control siRNA (siNeg), or mock treated (No siRNA). Lysates were immunoblotted with anti-HA and anti-β-actin antibodies. (**c**) Validation by cDNA rescue. MEFs stably expressing doxycycline-inducible *Renilla*-HA or *Abcf1*(rescue) cDNA were treated with siAbcf1(si-1) and doxycycline (0, 0.3, 3, 30 ug/mL). Cells were stimulated with ISD and CXCL10 production was measured by ELISA. * $P < 0.05$, *** $P < 0.001$ compared with siAbcf1-treated cells without doxycycline treatment.

(Rozenblatt-Rosen et al. 2012), suggesting that this DNA virus may derive a benefit from targeting this protein (**Fig. 1.2c**).

In the primary signaling network (**Fig. 2c**), known components such as *Sting*, *Tbk1*, and *Irf3* were strong hits in our screen; knockdown of these genes reduced ISD-stimulated CXCL10 production by more than 10-fold each (**Supplementary Table 1.3b,c**), validating our assay. Several screening hits that are hypothesized to interact with these proteins were also found to have phenotypes in our loss-of-function screen, including *Cdc37*, *Numa1*, and *Cyb5r3* (**Fig. 1.2c** and **Supplementary Table 1.3c**). Knockdown of *Cdc37* reduced ISD-stimulated CXCL10 production as strongly as the known components (**Supplementary Table 1.3c**). Consistent with this result, treatment of murine or human cells with celastrol, a small molecule inhibitor of the CDC37-HSP90 interaction (Gray et al. 2008; Zhang et al. 2008), potently reduced *Ifnb1* and CXCL10 induction (**Fig. 1.4a** and **Supplementary Fig. 1.4a**). CDC37 is a molecular co-chaperone that interacts with HSP90 to stabilize specific proteins, notably protein kinases (Gray et al. 2008), and is a putative interactor of TBK1 (Bouwmeester et al. 2004; Li et al. 2011). Thus, we tested whether CDC37 regulates TBK1 expression. Knockdown of *Cdc37* substantially reduced protein levels of TBK1 (**Fig. 1.4b**). As expected from this result, knockdown of *Cdc37* abrogated the phosphorylation of IRF3 (serine-396) – a hallmark of the ISD pathway's activation – following ISD stimulation (**Fig. 1.4b**). Chemical inhibition of HSP90 (by 17-DMAG) or of TBK1 (by BX795) had similar phenotypes as CDC37 inhibition (**Fig. 1.4a** and **Supplementary Fig. 1.4a**). Thus, targeting the members of this complex with small molecules can block the ISD response by potently inhibiting TBK1 protein stability or activity, a phenotype that may be applicable to the treatment of certain autoimmune disorders (see below).

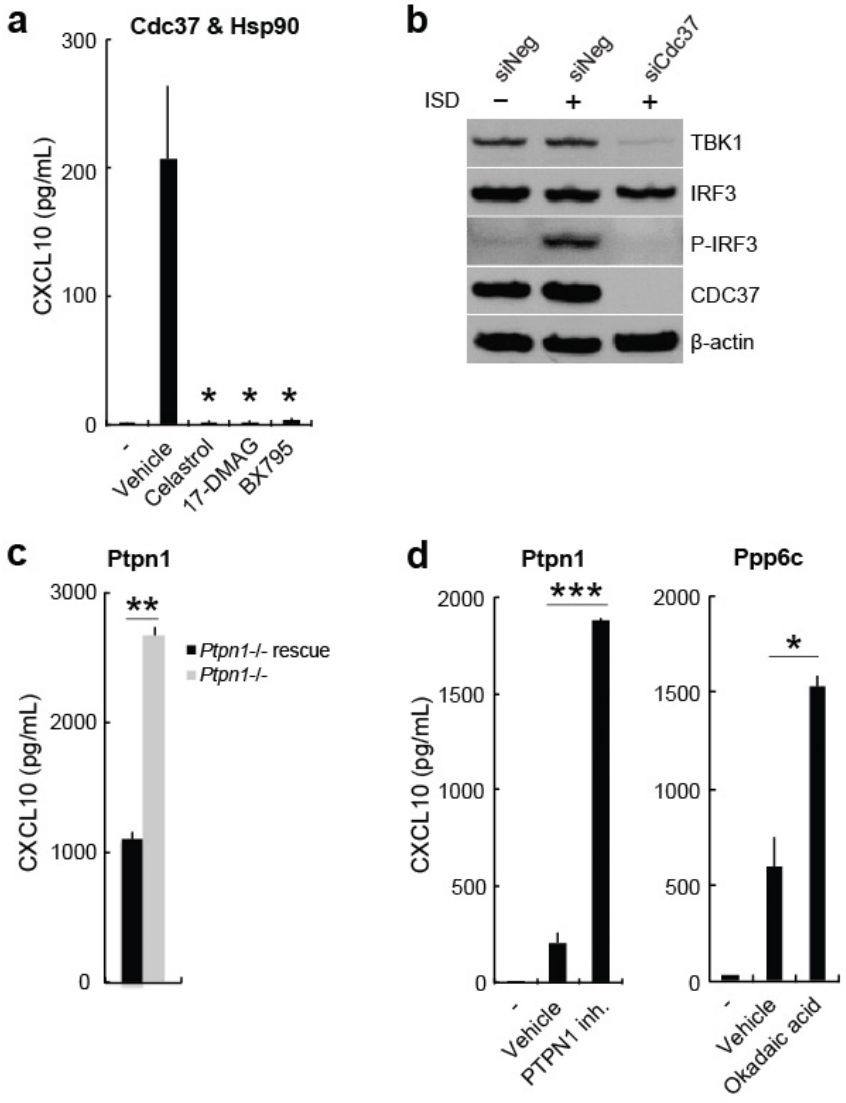


Figure 1.4. Targeting of screening hits by small molecule inhibitors. (**a**) Human monocyte-derived dendritic cells (MoDCs) treated with small molecule inhibitors were stimulated with 0.3 ug/mL DNA (HIV gag-100) for 24 h and CXCL10 production was measured by ELISA. Small molecules were used at: 500 nM celastrol, 75 nM 17-DMAG, and 500 nM BX795. * $P < 0.05$ compared with vehicle control. (**b**) MEFs treated with siRNAs were stimulated for 4 h with ISD, and lysates were immunoblotted with indicated antibodies. (**c**) *Ptpn1*-/- MEFs or *Ptpn1*-/- MEFs rescued with *Ptpn1* cDNA were stimulated with ISD, and CXCL10 production was measured by ELISA. ** $P < 0.01$. (**d**) Human MoDCs treated with 7.5 uM PTPN1 inhibitor or 1 nM okadaic acid were stimulated with 0.3 ug/mL DNA (HIV gag-100) and CXCL10 production was measured by ELISA. * $P < 0.05$, *** $P < 0.001$ compared with vehicle control.

Secondary signaling downstream of the IFN receptor (**Fig. 1.2c**) is also important in the ISD response, and we identified known (e.g. *Irf9* and *Stat1*) and candidate mediators (e.g. *Ptpn1* (Myers et al. 2001)) of the secondary signaling network (**Fig. 1.2c**, **Supplementary Fig. 1.4b**, and **Supplementary Table 1.3b,d**). Knockout of the protein tyrosine phosphatase, PTPN1, increased ISD-induced CXCL10 production 2.4-fold (**Fig. 1.4c** and **Supplementary Fig. 1.4c**), validating the screening phenotype. Consistent with this result, small molecule inhibition of PTPN1 increased CXCL10 production 9.1-fold in human monocyte-derived dendritic cells (MoDCs) stimulated with ISD (**Fig. 1.4d**).

Finally, we further tested hits in which the molecular interaction partners in the ISD pathway remain unclear. SP110 is an IFN-regulated nuclear body protein. We found that a natural deleterious mutation in *Sp110* (Pan et al. 2005) causes 1.5-3 fold decreased *Ifnb*1 induction in response to DNA (**Supplementary Fig. 1.4d**). The protein serine/threonine phosphatase, PPP6C, is a candidate interactor of IκB-ε (Bouwmeester et al. 2004), but its target in the ISD pathway is unknown. Consistent with the phenotype in our siRNA screen, we found that small molecule inhibition of PPP6C increased ISD-stimulated CXCL10 production in human MoDCs 2.6-fold (**Fig. 1.4d**). Thus, targeting the two inhibitory phosphatases, PTPN1 and PPP6C, by small molecules may serve as a way to enhance the immune response to certain DNA viruses and retroviruses (see below), and possibly to enhance the immunogenicity of DNA vaccines (Ishii et al. 2008; Ishikawa et al. 2009).

Quantitative proteomics analysis of the DNA-sensing network

We sought to further understand the DNA-sensing network (**Fig. 1.2c**) by determining the interaction partners of ABCF1, a protein that we found regulates the ISD response (**Fig. 1.2b** and

Supplementary Table 1.3a) and associates with DNA (**Fig. 1.1c** and **Supplementary Fig. 1.1d**). We again used an unbiased quantitative mass spectrometry-based approach, in which we precipitated stably-expressed ABCF1-HA in MEFs with anti-HA antibody and performed SILAC mass spectrometry (**Fig. 1.5a**). We identified 53 proteins with SILAC ratios that demonstrated co-precipitation with ABCF1 at a *P*-value < 0.01 (**Fig. 1.5a** and **Supplementary Table 1.5**). Three of the proteins that co-precipitated with ABCF1 – SET, HMGB2, and ANP32A – are members of the ER-associated SET complex (**Fig. 1.5b,c**), of which we had previously isolated HMGB2 and the DNA exonucleases TREX1 and APEX1 by DNA precipitation (**Fig. 1.1c,d**). None of these interactions were seen with STING-HA pulldown (**Supplementary Table 1.1**), indicating specificity for ABCF1. Immunofluorescence suggested that a subset of ABCF1 colocalizes with SET and the ER-marker calreticulin, further supporting the interaction between ABCF1 and the SET complex (**Fig. 1.5d** and **Supplementary Fig. 1.5a**). The ER co-localization is consistent with another report suggesting that ABCF1 is partially localized to the ER and partially localized to the cytosol (Paytubi et al. 2009).

The SET complex member, HMGB2 is hypothesized to function as a co-ligand for nucleic acid sensors though the precise role of HMGB2 remains unclear (Yanai et al. 2009). We observed that not only HMGB2, but also IFI204 – a predicted DNA sensor also known as IFI16 (Unterholzner et al. 2010) – are candidate interactors of ABCF1 (**Fig. 1.5b,c**). Consistent with the reported functions of *Hmgb2* and *Ifi16* (Yanai et al. 2009; Unterholzner et al. 2010), knockdown of *Abcf1* suppressed TBK1 and IRF3 phosphorylation in MEFs stimulated with ISD (**Fig. 1.5e** and **Supplementary Fig. 1.5b,c**).

While *Abcf1* knockdown significantly reduced *Ifnb1* and ISG induction following dsDNA (HIV gag-100 sequence) stimulation or HSV-1 *d109* infection (**Fig. 1.5f,g** and **Supplementary**

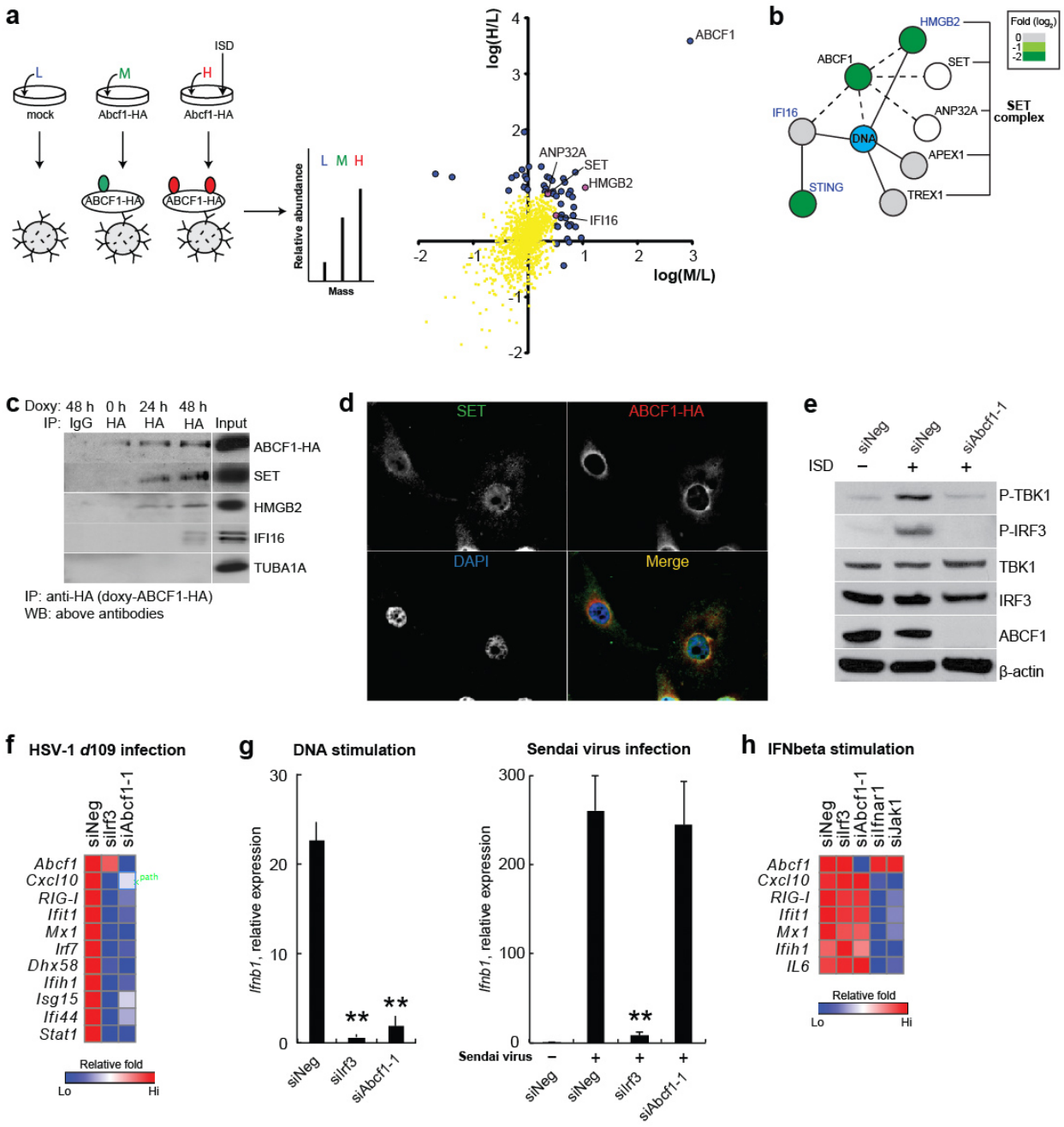


Figure 1.5. Identification of components of DNA sensing complex. (**a**) ABCF1 interactors. MEFs stably expressing *Abcf1*-HA or mock-transduced MEFs were stimulated with ISD or left unstimulated. Lysates were precipitated with anti-HA antibody. X-axis (M/L) and y-axis (H/L) correspond to co-precipitation with ABCF1-HA in the unstimulated and stimulated states, respectively. Blue/purple circles, hits with *P*-value < 0.01; yellow squares, non-hit. (**b**) Interaction among proteins in DNA sensing network. Legend is same as in Figure 2c. (**c**) MEFs stably expressing doxycycline-inducible *Abcf1*-HA were treated with doxycycline for 0, 24, or 48 h. Lysates were precipitated with anti-HA antibody or IgG control. Precipitates were immunoblotted with indicated antibodies. (**d**) Immunofluorescent microscopy of MEFs expressing HA-tagged *Abcf1*, stained for DAPI, HA, and SET. Merge of images is shown. (**e**) MEFs treated with indicated siRNAs were stimulated for 4 h with ISD, and lysates were immunoblotted with indicated antibodies. (**f**) MEFs treated with indicated siRNAs were infected with

HSV-1 *d*109 for 6 h, and induction levels of *Cxcl10* mRNA were determined by qRT-PCR. Data presented as mean and s.d. (n = 4). (**g**) *Trex1*-/- MEFs treated with indicated siRNAs were stimulated with 4 ug/mL DNA (HIV gag-100) for 5.5 h or infected with Sendai virus for 6 h, and induction levels of *Ifnb1* mRNA were determined by qRT-PCR. Data presented as mean and s.d. (n = 3). In (f) and (g), ** $P < 0.01$, ** $P < 0.001$ compared with cells treated with control siRNA. (**h**) *Trex1*-/- MEFs treated with indicated siRNAs were stimulated with 300 U/mL IFNβ for 8 h, and expression levels of *Abcf1* and ISG mRNA were determined by qRT-PCR. Data are averages of triplicate wells.

Fig. 1.5d,e), *Abcf1* knockdown did not have a significant effect on IFN or ISG induction by Sendai virus (which stimulates the RIG-I pathway) or by recombinant IFNβ itself (**Fig. 1.5g,h** and **Supplementary Fig. 1.5f**). Taken together, these results support the hypothesis that ABCF1 interacts with HMGB2, IFI16, and the SET complex and is a critical node in the DNA sensing network.

Innate immune response to retroviral infection in *TREX1* mutant cells

We further examined whether our findings are relevant to *Trex1*-dependent autoimmunity and retroviral infection. While in culture *Trex1*-/- cells do not spontaneously produce type I IFNs or ISGs, retroviral infection of these cells induces IFN and ISG production in the absence of *Trex1* (**Supplementary Fig. 1.6a** and ref. (Yan et al. 2010)). We knocked down our major screening hits in *Trex1*-/- MEFs and infected the cells with an HIV-based retrovirus. Knockdown of 4 of these genes (i.e. *Ptpn1*, *Tiparp*, *Mdp1*, and *Ppp6c*) significantly enhanced the ability of retroviral infection to induce IFN and ISG production ($P < 0.05$), while knockdown of 10 of these genes (including *Abcf1* and *Cdc37*) as well as 4 known signal transduction components (i.e. *Trim56*, *Sting*, *Tbk1*, and *Irf3*) significantly abrogated the innate immune response ($P < 0.05$; **Fig. 1.6a-c** and **Supplementary Fig. 1.6b**).

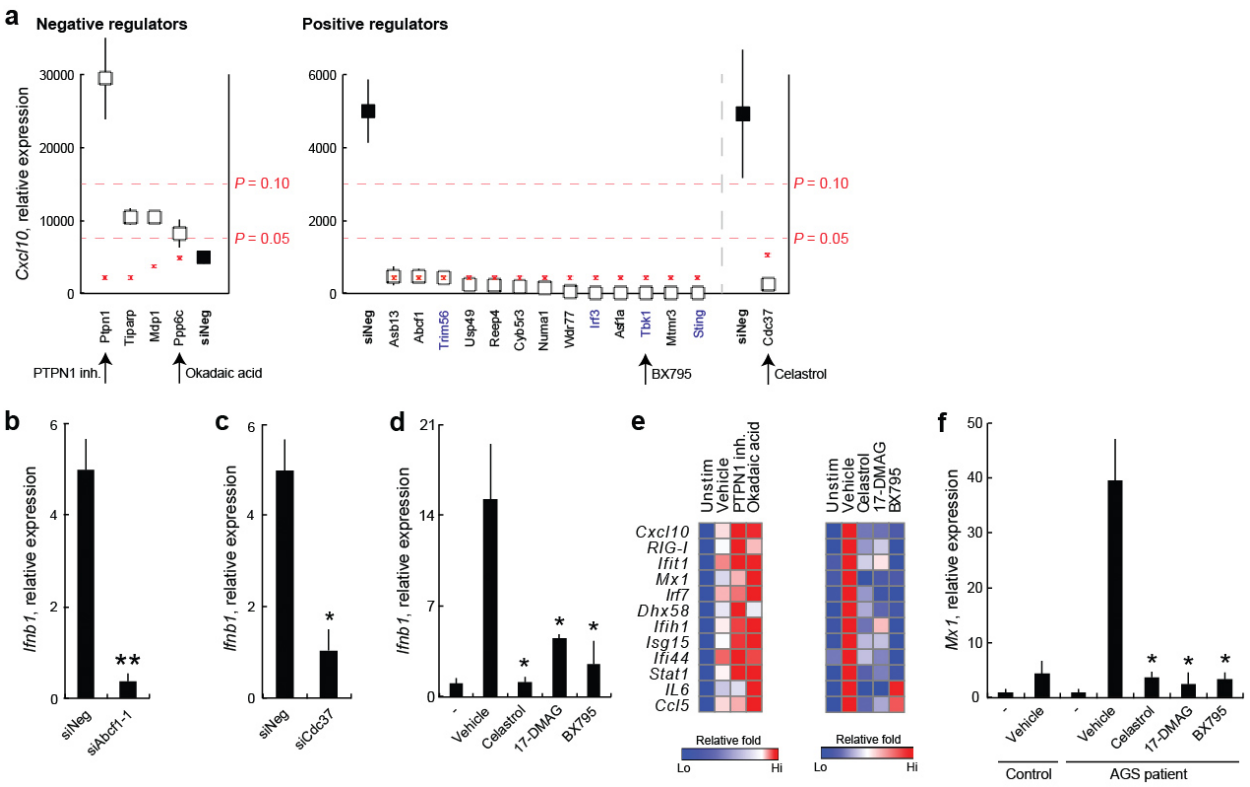


Figure 1.6. Inhibition of identified regulators by RNAi or small molecules modulates the innate immune response to retroviral infection. (**a**) *Trex1*-/- MEFs treated with indicated siRNAs were infected with retrovirus for 21.5 h, and induction levels of *Cxcl10* mRNA were determined by qRT-PCR. *P*-values compared with control siRNA (siNeg)-treated cells (after Benjamini-Hochberg correction for multiple testing) are marked as asterisks in red with $P = 0.05$ and $P = 0.10$ indicated as dotted lines. Data presented as mean and s.d. (n = 3). Known ISD pathway genes are colored blue. Genes whose protein products can be inhibited by existing small molecule inhibitors are marked with arrows. (**b,c**) *Trex1*-/- MEFs treated with indicated siRNAs were infected with retrovirus, and induction levels of *Ifnb1* mRNA were determined by qRT-PCR. Data presented as mean and s.d. (n = 3). (**d,e**) *Trex1*-/- MEFs treated with small molecule inhibitors were infected with retrovirus, and induction levels of *Ifnb1* or a panel of ISGs was determined by qRT-PCR. Data are averages of duplicate wells. Small molecules were used at: 400 nM celastrol, 750 nM 17-DMAG, 500 nM BX795, 30 uM PTPN1 inhibitor, and 10 nM okadaic acid. (**f**) *TREX1* mutant (AGS patient 1: R114H/D201ins; AGS patient 2: R114H/R114H) human fibroblasts and healthy control fibroblasts were treated with vehicle alone or small molecule inhibitors and infected with retrovirus. Induction levels of the ISG, *MX1*, was determined by qRT-PCR. Data are averages of triplicate wells. Small molecules were used at: 500 nM celastrol, 100 nM 17-DMAG, and 500 nM BX795. * $P <$ 0.05 compared with respective cells that were treated with vehicle control and infected with retrovirus.

Chemical inhibition of CDC37, HSP90, or TBK1 potently abrogated retroviral infection-induced *Ifnb1* induction in *Trex1*-/- MEFs (**Fig. 1.6d,e** and **Supplementary Fig. 1.6c**); in contrast, chemical inhibition of PTPN1 or PPP6C increased ISG induction in response to retroviral infection in a dose-dependent manner (**Fig. 1.6e** and **Supplementary Fig. 1.6c**). We then tested several of these small molecules in human fibroblasts derived from two healthy individuals and from two patients with Aicardi-Goutières syndrome that had disease-causative mutations in TREX1 (R114H/D201ins in one patient, and R114H/R114H in the second patient) (Crow et al. 2006). The R114H variant is also found in patients with SLE (Lee-Kirsch et al. 2007b). The small molecules targeting CDC37, HSP90, and TBK1 abrogated the induction of ISGs (**Fig. 1.6f** and **Supplementary Fig. 1.6d**). Thus, celastrol, 17-DMAG, and BX795 may be therapeutic leads in *TREX1*-dependent autoimmune disorders, while PTPN1 inhibitor and okadaic acid may exacerbate the autoimmune phenotype but enhance the innate immune response to retroviral infection.

DISCUSSION

We describe here the integration of complementary genomic and proteomic datasets to identify new components and physical interactions in the ISD signaling network. We generated and used DNA-protein interaction, protein-protein interaction, and loss-of-function screening datasets to identify new ISD pathway components; validated several of the newly identified components; demonstrated that a subset also function in the response to retroviral infection in *Trex1*-/- cells; and showed that small molecule inhibitors of several of these components can modulate the innate immune response to dsDNA and retroviral infection.

In our DNA precipitation experiment, we identified 184 candidate cytoplasmic DNA-interacting proteins that encompass most of the published components of DNA-sensing pathways (including ISD, RNA polymerase III, AIM2 inflammasome, and AGS proteins), with several exceptions (e.g. AIM2 and DDX41 which may be specific to cells of the myeloid lineage (Schroder and Tschopp 2010; Zhang et al. 2011b)). After screening the candidates for a potential role in the ISD response, we identified ABCF1 as a cytoplasmic protein that associates with dsDNA, IFI16, and HMGB2, and regulates the interferon response to transfected dsDNA and retroviral infection. These results implicate ABCF1 as a key component of the ISD pathway. It is still unclear whether ABCF1 interacts with DNA directly, or whether the interaction is indirect.

Our experiments also identified SET complex members (SET, ANP32A, and HMGB2) as ABCF1 interactors. The SET complex contains three DNA nucleases (TREX1, APEX1, and NME1); the chromatin-modifying proteins SET and ANP32A; and HMGB2, which functions as a co-receptor for nucleic acid receptors among other roles (Chowdhury and Lieberman 2008). The observed interactions among dsDNA, ABCF1, HMGB2, and other SET complex members suggest that early steps in DNA recognition may occur at the ER-localized SET complex.

Consistent with this hypothesis, the complex member TREX1 can prevent HIV-1 DNA detection, and its absence results in accumulation of retroelement DNA at the ER which drives an ISD response (Yang et al. 2007; Stetson et al. 2008; Yan et al. 2010). Furthermore, the complex members SET and NME1 also detect HIV-1 DNA, and in turn regulate HIV-1 infectivity[27]. A recent model suggests that the SET complex may recognize viral DNA as damaged DNA, specifically via its base excision repair (BER) activity and/or its distorted structure (e.g. HMGB2) (Yan et al. 2011). Consistent with this model, we find that our ISD interactors include the SET and BER complex member, APEX1, as well as nearly the entire BER complex (e.g. PARP1, PARP2, POLB, LIG3, XRCC1, FEN1, and PCNA). Overall, these results suggest that the SET complex plays a central role in DNA sensing and forms a coordinated system for detecting, modifying, and degrading viral or retroelement DNA.

We tested the impact of small molecule inhibitors on the DNA-sensing response and found that inhibition of PTPN1, PPP6C, CDC37, HSP90, or TBK1 modulates the innate immune response to cytosolic DNA in human dendritic cells and to retroviral infection in *TREX1* mutant cells. Several of these small molecules have already been tested for other conditions in mice and humans. For example, 17-DMAG is being explored for the treatment of autoimmune disease and various cancers (Shimp et al. 2012), and extracts of the medicinal plant *Tripterygium wilfordii*, from which the natural product celastrol is derived, have been used as an anti-inflammatory (Corson and Crews 2007; Goldbach-Mansky et al. 2009). While current treatments for *Trex1*-dependent autoimmune disorders (including Aicardi-Goutières syndrome, familial chilblain lupus, and SLE) do not target the cause of these diseases, small molecules like celastrol, 17-DMAG, and BX795 that inhibit the ISD response may represent new therapeutics for this class of disorders.

MATERIALS AND METHODS

Cells, viruses, and reagents. *p53*-/- MEFs were a gift from D.M. Sabatini and D.J. Kwiatkowski (Zhang et al. 2003). *Trex1*-/- MEFs and C57BL/6 wild type (wt) control MEFs were a gift from T. Lindahl. *Ptpn1*-/- MEFs and *Ptpn1*-/- MEFs rescued with *Ptpn1* cDNA were a gift from B.G. Neel (Klaman et al. 2000). Human AGS patient (*TREX1* R114H/D201ins and *TREX1* R114H/R114H) fibroblasts and healthy control cells were a gift from Y.J. Crow. 293T cells were obtained from ATCC. Primary murine lung fibroblasts were derived from lung tissue of 4-8 wk old female C57BL/6 mice as previously described (Tager et al. 2004). cDCs were prepared from B6.C3H-sst1 (*Sp110* LoF) as previously described (Pan et al. 2005; Amit et al. 2009). Cells were maintained in DMEM (Mediatech) supplemented with 10% FBS (Sigma). Human monocytes were isolated by negative selection (Life Technologies) from peripheral blood mononuclear cells, and differentiated into dendritic cells by a seven day culture in GM-CSF (R&D) and IL-4 (R&D) in RPMI (Life Technologies) supplemented with 10% FBS (Life Technologies). Sendai virus was obtained from ATCC and used at an MOI of 1. HSV-1 *d*109 (Samaniego et al. 1998) was obtained as a gift from N.A. DeLuca and was used at an MOI of 10. Self-inactivating minimal HIV-1 virus was produced in 293T cells using the vector pLX301 (TRC, Broad Institute), the packaging construct psPAX2, and the envelope plasmid pCMV-VSVG. Interferon stimulatory DNA (ISD), HIV gag-100, and HSV60 dsDNA were annealed from oligonucleotides (IDT) as described previously (Stetson and Medzhitov 2006; Unterholzner et al. 2010; Yan et al. 2010). Sequences are listed in Supplementary Table 1.5a. *In vitro* transcribed RNA was synthesized as described previously(Hornung et al. 2006). Nucleic acids were mixed with Lipofectamine LTX (Life Technologies) at a ratio of 1:3 (wt/vol) in Opti-MEM (Life Technologies), and added to cells at a final concentration of 1 ug/mL (DNA) or 0.1 ug/mL (RNA) unless otherwise indicated.

Recombinant IFNβ was obtained from PBL InterferonSource; murine CXCL10 ELISA kit was obtained from R&D; NE-PER from Pierce; Coomassie blue (SimplyBlue SafeStain) from Life Technologies; EXPRESS35S Protein Labeling Mix from Perkin Elmer. Antibodies were obtained from the following sources: anti-P-TBK1 Ser172 (5483; Cell Signaling), anti-P-IRF3 Ser396 (4947; Cell Signaling), anti-TBK1 (3504; Cell Signaling), anti-IRF3 (4302; Cell Signaling), anti-CDC37 (4793; Cell Signaling), anti-ABCF1 (SAB2106638, Sigma), anti-HMGB2 antibody (ab67282, Abcam), anti-SET (sc25564, Santa Cruz), anti-IFI204 (SAB2105265, Sigma), anti-α-tubulin (T5168, Sigma), anti-β-actin (ab6276, Abcam), anti-HA (High Affinity 3F10; Roche), anti-SMARCB1 (H-300; Santa Cruz Biotechnology), anti-calreticulin (ab14234; Abcam), and rat IgG control (Jackson Laboratories). PTPN1 inhibitor (CAS 765317-72-4), okadaic acid, and celastrol were obtained from Millipore. BX795 and 17-(Dimethylaminoethylamino)-17-demethoxygeldanamycin (17-DMAG) were obtained from Invivogen.

Identification of DNA-interacting proteins by SILAC mass spectrometry. *p53*-/- MEFs were grown for six cell doublings in DMEM depleted of L-arginine and L-lysine (Caisson Labs Inc.) and supplemented with 10% dialyzed FBS (Sigma) and either light (L), medium (M), or heavy (H) isotope-labeled amino acids . The L and H cells were stimulated with 1000 U/mL IFNβ for 18 h. The cells were pelleted and incubated in hypotonic lysis buffer (10 mM HEPES pH 7.4, 10 mM KCl, 1 mM EDTA) containing protease inhibitors (Roche) for 10 min on ice followed by lysis for 1 min in hypotonic lysis buffer supplemented with 0.3% Triton X-100. Nuclei and insoluble proteins were removed by centrifugation. 11 mg of the H and L lysates were mixed with a 1:1 mix of biotinylated ISD and ISD with a tetraethylene glycol arm between the biotin

and the nucleic acid (IDT), and then split into two samples. No ISD was added to 11 mg of sample M. Equal volumes of streptavidin beads (Ultralink; Pierce) were added to all three samples, and samples were rotated for 2.5 h at 4°C. Beads were pelleted and washed extensively with wash buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 1 mM EDTA, 0.75% NP-40, 0.175% sodium deoxycholate). The three samples were mixed, cysteines were reduced by DTT and alkylated with iodoacetamide, and proteins were eluted by heating in SDS sample buffer (Life Technologies) for 10 min before separation on a 4-12% gradient gel (NuPAGE; Life Technologies). The resolved proteins were divided into 13 fractions and subjected to proteolysis with trypsin, and peptides were extracted from the gel as described (Shevchenko et al. 2006). Peptide extracts were cleaned up offline with C18 StageTips (Rappsilber et al. 2007) prior to 90 min nanoESI-LCMS analyses with a gradient of 3%-35% acetonitrile/ 0.1% formic acid. Protein and peptide identification and quantification was performed with MaxQuant (v. 1.0.12.31) (Cox et al. 2011) using the IPI mouse v3.52 as the search database. Search parameters specified trypsin cleavage with 2 missed cleavages, peptide mass tolerance of 6 ppm and fragment mass tolerance of 0.5 Da; carbamidomethylated cysteines and variable modifications of oxidized methionine, acetylation of protein N-termini and sample-specific modifications of Arg-0,6,10 and Lys-0,4,8 for SILAC triple labeling. Protein ratios were medians of ratios from at least two quantified peptides. To identify significant proteins, *P*-values were calculated via Gaussian modeling of the log(H/M) data, and a significance threshold of $P < 1x10^{-4}$ was used.

Identification of STING- and ABCF1-interacting proteins by SILAC mass spectrometry. Light (L), medium (M), and heavy (H) isotope-labeled *p53*-/- MEFs mock infected (L) or stably expressing *Sting*-HA (H) or *Abcf1*-HA (M and H) were stimulated with 1000 U/mL IFNβ for 18

h. H cells were transfected with ISD for 2.5 h. Cells were lysed in 50 mM Tris-HCl pH 7.8, 150 mM NaCl, 1 mM EDTA, 0.2% NP-40, and insoluble proteins were removed by centrifugation. 18 mg of each lysate was mixed with 1 ug/mL anti-HA antibody and Protein G beads (Pierce), and rotated for 2.5 h at 4°C. Beads were pelleted and washed extensively with wash buffer (above). The isotope-labeled samples were mixed (with *Sting*-HA and *Abcf1*-HA samples handled separately), cysteines were reduced by DTT and alkylated with iodoacetamide, and proteins were eluted by heating in SDS sample buffer for 10 min before separation on a 4-12% gradient gel. The whole gel lane was cut into 8 slices, proteins were digested inside the gel with trypsin, and peptides were extracted from the gel. Extracted peptides were purified with StageTips and analyzed with a 100 min acquisition method on a Thermo EASY-nLC 1000 UHPLC coupled to a Q Exactive mass spectrometer. Raw data files were processed in MaxQuant (v. 1.2.2.5) using the IPI mouse v3.68 as the search database. All proteins were identified with at least 2 or more unique peptides and quantified with 3 or more ratios. SILAC ratios were normalized over the median of all protein ratios to correct for sample losses between parallel immunoprecipitation steps. To identify significant interactions, *P*-values were calculated via Gaussian modeling of the log(M/L) and log(H/L) data, and a significance threshold of $P < 0.01$ was used.

RNA interference screen. 750 *p53*-/- MEFs per well were seeded in 96-well plates in 60% DMEM and 40% Opti-MEM. 25 nM siRNA was complexed with 0.5 uL Lipofectamine RNAiMax (Life Technologies) in Opti-MEM, incubated for 12 min at 22°C, and added to the wells. 72 h later, cells were transfected with ISD. 26 h later, supernatants were collected and CXCL10 was quantified by ELISA. Cell viability was estimated by the CellTiter-Glo

Luminescent Cell Viability Assay (Promega); CellTiter-Glo values below 3.75e5 were considered toxic. Dharmacon siGENOME SMARTpools from Harvard ICCB were used for screening. ON-TARGETplus Non-targeting Pool was used as negative control (siNeg). Individual siRNAs were from Dharmacon, Life Technologies, Qiagen, and Sigma. siRNA sequences are listed in Supplementary Table 1.5b.

Plasmid construction. To create the tet-on lentiviral vector (pCW57d-P2AR), the pLKO.1 (Moffat et al. 2006) vector was modified as follows: the U6 shRNA cassette was removed from LKO.1 and the TRE with a MCS was inserted upstream of the PGK promoter; the rtTA was cloned 3' of the puroR (with a 2A multicistronic cleavage site between these two genes) along with a WPRE. To create the *Abcf1*(rescue) construct, silent mutations were made in the siAbcf1(si-1) targeting site using overlap extension PCR; the rescue cDNA was then cloned into pCW57d-P2AR. The *Irf3*(rescue) construct was similarly cloned, with silent mutations made in the siIrf3 targeting site. To generate the HA-tagged *Abcf1* and HA-tagged *Sting* expression vectors, an HA tag was added to the C-terminus of the cDNAs during PCR, and the constructs were cloned into pLX301. Primer sequences are listed in Supplementary Table 1.5c.

cDNA rescue. *p53*-/- MEFs stably expressing cDNA in the pCW57d-P2AR vector were subjected to siRNA. 72 h later, doxycycline (Sigma) at 0, 0.3, 3, or 30 ug/mL was added to the cells, and cells were stimulated with ISD. 26 h later, supernatants were collected and CXCL10 was quantified by ELISA.

Quantitative RT-PCR. Total RNA was prepared from cells using the RNeasy Mini kit (Qiagen). cDNA was synthesized using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Real time qPCR was performed using SYBR Green (Roche) and the LightCycler 480 system (Roche) according to instructions provided by the manufacturer. Relative amounts of mRNA were normalized to *Gapdh* levels in each sample. The primers used for qPCR are listed in Supplementary Table 1.5d.

Co-immunoprecipitation assays. *p53*-/- MEFs stably expressing doxycycline-inducible *Abcf1*-HA were treated with 3 ug/mL doxycycline for 0, 24, or 48 hours, and then lysed in 10 mM Tris-HCl, 2 mM EDTA, 0.4% NP-40 with complete EDTA-free protease inhibitors (Roche). Anti-HA antibody was crosslinked to Protein G beads (Roche) at a concentration of 1 ug antibody per 20 uL beads using dimethyl pimelimidate dihydrochloride (Sigma). Cleared supernatants were incubated with the antibody-bound beads, and rotated overnight at 4°C. Rat IgG control bound to Protein G were used as IP control with the 48 h doxy-treated lysates. The beads were washed extensively with wash buffer (10 mM Tris-HCl, 2 mM EDTA, 1% NP-40). Immunoprecipitates were eluted with 100 uL 2.5 M Glycine (pH 3) and immediately buffered to pH 7.5 by adding 25 uL of 2 M Tris. Samples were boiled in reducing Laemmli buffer for 10 min, separated by SDS-PAGE, and immunoblotted using anti-HA, anti-SET, and anti-HMGB2 antibodies.

Immunofluorescence assays. $5x10^4$ *p53*-/- MEFs stably expressing doxycycline-inducible *Abcf1*-HA were grown overnight on glass coverslips in standard DMEM with 10% FBS. *Abcf1*-HA expression was induced by adding 3 ug/mL doxycycline for 48 h. Cells were fixed using 4% PFA for 15 min and permeabilized using PBS + 0.2% Triton X-100. Cells were then treated with

respective primary antibodies at a concentration of 1/200 for 1 h, followed by treatment with respective fluorophore labeled secondary antibodies. The cells were mounted on glass sides using DAPI-containing VECTASHIELD (Vector Laboratories). The cells were visualized using 503 Platform from Intelligent Imaging Innovations, under 40X and 63X oil immersion. 8 Z-stacks were taken per image at 1 um per step and the image were deconvolved using nearest neighbor algorithm. The images were processed using Slidebook version 5.

DNA microarray analysis. (i) 293T cells were stimulated with 1000 U/mL recombinant IFNβ followed by lysis 6 h later in RLT buffer (Qiagen). Total RNA was isolated using the RNeasy Mini kit (Qiagen). Genome-wide gene expression profiling was obtained by hybridizing the RNA to the Affymetrix Human U133 Plus 2.0 array. The cDNA synthesis, labeling, and subsequent hybridization to the microarrays were performed at the Molecular Profiling Laboratory, MGH Center for Cancer Research. (ii) *p53*-/- MEFs were treated with siRNAs for 72 h and then transfected with ISD. 6 h later, cells were lysed in RLT buffer. Each sample was performed in biological duplicates. Total RNA was isolated using the RNeasy Mini kit. RNA quantification and quality was assessed using Agilent 2100 Bioanalyzer (Agilent Technologies). Genome-wide gene expression profiling was obtained by hybridizing the RNA to the Affymetrix GeneChip® Mouse Gene 1.0 ST Array. The cDNA synthesis, labeling, and subsequent hybridization to the microarrays were performed by the company Expression Analysis (Durham, NC).

Curation of microarray data. Genes were curated as follows: (i) Top 340 upregulated genes (>6.1-fold upregulation after stimulation of MEFs with poly(dA-dT)-poly(dT-dA) for 4 h) from

GDS1773 (Ishii et al. 2006); (ii) Top 200 upregulated genes (>3.9-fold upregulation after stimulation of NIH3T3 cells with recombinant IFNb for 4 h), and top 300 upregulated genes (>6.25-fold upregulation after stimulation of L929 cells with recombinant IFNb for 4 h) from GSE14413 (Burckstummer et al. 2009). (iii) Top 150 upregulated genes (>1.3-fold upregulation after stimulation of 293T cells with recombinant IFNβ for 8 h) from above-described arrays.

Network analysis. Network analysis was carried out using PPI data from Ingenuity, the STRING database (http://string.embl.de), and PPIs found experimentally in published studies (Bouwmeester et al. 2004; Cristea et al. 2010; Li et al. 2011; Pichlmair et al. 2012; Rozenblatt-Rosen et al. 2012) and in above-described SILAC experiments.

Statistics. Statistical significance was determined by Student's t-test.

CHAPTER 2: A GENETICAL GENOMICS APPROACH TO UNDERSTANDING THE REGULATORY MECHANISMS OF HOST-PATHOGEN INTERACTIONS IN HUMANS

Mark N. Lee*, Chun Ye*, Alexandra-Chloé Villani, Towfique Raj, Weibo Li, Selina H. Imboywa, Portia I. Chipendo, Khadir Raddassi, Michelle H. Lee, Irene Wood, Cristin McCabe, Barbara E. Stranger, Christophe O. Benoist, Philip De Jager, Aviv Regev, Nir Hacohen

AUTHOR CONTRIBUTIONS

Lee MN is responsible for helping to conceive the design of the project, Ficolling some of the blood samples, designing the high-throughput cellular assays, optimizing stimulation conditions, preparing the monocyte-derived dendritic cells (MoDCs) for the microarray and Nanostring experiments, helping to choose the Nanostring codeset, running the Nanostring samples, and writing the manuscript.

Chun Ye is responsible for computationally analyzing the microarray and Nanostring data, helping to choose the Nanostring codeset, performing the eQTL associations, and helping to write the manuscript.

Villani AC is responsible for helping to prepare the MoDCs, helping to coordinate the microarray samples, and helping to choose the Nanostring codeset.

Raj T is responsible for analyzing the genotyping data.

Li W is responsible for Ficolling some of the blood samples, helping to prepare the MoDCs, and helping to run the Nanostring samples.

Imboywa SH is responsible for Ficolling most of the blood samples.

Chipendo PI is responsible for Ficolling some of the blood samples.

Raddassi K is responsible for Ficolling some of the blood samples.

Lee MH is responsible for helping to coordinate donor recruitment and blood collection.

Wood I is responsible for helping to coordinate donor recruitment and blood collection.

McCabe C is responsible for coordinating the DNA genotyping.

Stranger BE is responsible for helping to supervise the computational analysis and DNA genotyping.

Benoist CO is responsible for helping to supervise the project.

Jager PD is responsible for supervising the donor recruitment, blood collection, Ficolling, and DNA genotyping.

Regev A is responsible for helping to conceive the design of the project and supervising the computational analyses.

Hacohen N is responsible for conceiving the design of the project and supervising the entire project.

ACKNOWLEDGMENTS

We are grateful to Q. Sievers and C. Wu for advice about preparing primary cells. This work was supported by an NIH GO grant. M.N.L. is supported by an NIH MSTP fellowship.

ABSTRACT

While naturally occurring genetic variants are known to modulate gene expression (e.g. eQTL), limited studies have evaluated the influence of these variants on pathogen responses using relevant human primary immune cells. Results from such gene-by-environment studies will be crucial to further understand disease etiology of a variety of disorders including infectious and autoimmune/autoinflammatory diseases. Here we present a systematic study evaluating variation of innate immune responses across a healthy cohort of individuals. We measured gene expression of primary monocyte-derived dendritic cells (MoDCs) activated through bacterial and viral receptors that mediate innate immune responses to pathogen components. We identified natural genetic variants that are associated with the variation in these responses, with some of the variants already known to be associated with disease susceptibility.

INTRODUCTION

The human innate immune system is of central importance to the early containment of infection. When receptors of innate immunity recognize molecular patterns on pathogens, they initiate an immediate immune response by inducing the expression of cytokines and other host defense genes (Takeuchi and Akira 2010). Altered expression or function of the receptors, the signaling molecules activated downstream of the receptors, or the induced cytokines themselves can alter individuals' propensity to acquire infectious or autoimmune diseases.

The Toll-like receptor (TLR) and RIG-I-like receptor (RLR) pathways of the innate immune system play key roles in detecting infection by various microorganisms, including many viruses and bacteria (Akira et al. 2006). Upon activation, TLRs signal through adaptors such as MYD88 and TRIF to induce the expression of pro-inflammatory and anti-viral genes, while RLRs signal through the MAVS and/or STING adaptors to induce a similar set of host defense genes including the anti-viral cytokine, IFNβ (Akira and Takeda 2004; Barbalat et al. 2011). IFNβ stimulates the heterodimeric type I interferon receptor composed of IFNAR1 and IFNAR2, which activates JAK and STAT adaptors to induce the expression of interferon-stimulated genes (ISGs) (Stark et al. 1998).

Genetic variation in such innate immune components has been shown to increase susceptibility to infectious and autoimmune diseases. For example, genetic studies of primary immunodeficiencies (PIDs) in humans have revealed that patients with Mendelian mutations in *MYD88* suffer from recurrent infection by *Streptococcus pneumonia* leading to meningitis or septicemia, as well as recurrent infection by *Staphylococcus* and *Pseudomonas* (von Bernuth et al. 2008). Mendelian mutations in *TLR3* or *TRIF* increase propensity to herpes simplex encephalitis (HSE) (Zhang et al. 2007; Guo et al. 2011; Sancho-Shimizu et al. 2011), while

mutations in *STAT1* or the JAK kinase family member, *TYK2*, lead to recurrent infection by viruses such as varicella zoster and herpes simplex virus (Dupuis et al. 2003; Casanova et al. 2012; Kilic et al. 2012). Mendelian mutations in the TLR-stimulated regulatory cytokine, *IL10*, cause infant colitis (Glocker et al. 2010), and variants in *IL10* have also been associated with type 1 diabetes mellitus (Barrett et al. 2009). Through genome-wide association studies (GWAS) common genetic variants in the type III interferon *IL28B*, which can be induced by viral stimulation of TLR and RLR pathway (Coccia et al. 2004), have been associated with the ability to clear hepatitis C viral infection (Ge et al. 2009; Suppiah et al. 2009; Tanaka et al. 2009; Thomas et al. 2009). Common variants in the IFN-stimulated gene, *IFITM3*, have been associated with morbidity following influenza infection (Everitt et al. 2012).

A catalog of functional variants affecting the human innate immune system would aid in the search for new markers of risk for infectious, autoimmune, or other inflammatory diseases, and may help uncover the molecular mechanism behind disease variants of unknown function (Cookson et al. 2009; Montgomery and Dermitzakis 2011; Dermitzakis 2012). Several studies have measured variation in gene expression in immune cell types such as lymphoblastoid cell lines (LCLs), primary B cells, and primary monocytes; and associated this variation with genotype data (Stranger et al. 2007; Zeller et al. 2010; Barreiro et al. 2012; Fairfax et al. 2012; Stranger et al. 2012). Such studies have identified genetic variants, called expression quantitative trait loci (eQTL), that associate with the expression levels of specific genes with statistical robustness (Schadt et al. 2003; Cookson et al. 2009). More recently, studies have identified genetic variants that associate with the change in expression in response to stimuli such as radiation and treatment with oxidized lipids (Smirnov et al. 2009; Romanoski et al. 2010). These eQTLs represent gene-by-environment interactions that are not captured by baseline expression

studies. Limited studies have systematically measured the capacity of individuals to mount innate immune responses and assessed whether underlying genetic variation explains variation in these responses.

Here we systematically assessed variation in the innate immune responses of primary immune cells of healthy humans. We profiled gene expression of cells: (i) at baseline; (ii) stimulated with the purified bacterial ligand, lipopolysaccharide (LPS) from *E. coli*; (iii) infected with the virus, influenza A ΔNS1; or (iv) stimulated with the cytokine, IFNβ. LPS stimulates the canonical TLR, TLR4; viral RNA from influenza stimulates the key viral sensors, RIG-I and TLR3, and deletion of the viral NS1 gene prevents inhibition of the RIG-I pathway (Kato et al. 2006; Mibayashi et al. 2007; Loo et al. 2008; Gack et al. 2009); and IFNβ stimulates the type I interferon receptor (IFNAR1 and IFNAR2) which controls the expression of a large number of anti-viral genes (Uze et al. 1990; Novick et al. 1994). Because prior studies have suggested that many eQTL are cell type specific (Dimas et al. 2009), we measured gene expression in primary monocyte-derived dendritic cells (MoDCs) since DCs play a central role in detecting infection and link innate immune activation to activation of the adaptive immune system (Mellman and Steinman 2001). We assessed the robustness of these expression phenotypes by conducting a serial replication study and identified innate immune responses that are highly reproducible. We then confirmed that genetic variation underlies some of these phenotypes.

RESULTS

Establishment of a system to quantitatively measure innate immune responses in MoDCs

We first established a robust, high-throughput system to quantitatively assess innate immune responses in primary MoDCs (**Fig. 2.1a,b**). To minimize batching of samples, we developed a method to isolate $CD14^{+}CD16^{lo}$ monocytes from peripheral blood mononuclear cells (PBMCs) in 96-well plates that yielded a median of 93% $CD14^{+}$ cells and 98% $CD16^{lo}$ cells.

Next, we selected conditions to stimulate and infect the MoDCs. We stimulated MoDCs pooled from 5 donors with a dose curve of LPS or influenza ΔNS1 (**Supplementary Fig. 2.1a**). We read out gene expression using the Nanostring nCounter system (Geiss et al. 2008), which is a hybridization-based method to digitally count transcripts. For each stimulus, we selected a dose near the average saturation points of the dose-response curves. The doses we selected (e.g. 15 ng/mL LPS) are consistent with standard doses used in the field. We then performed a time course with these stimuli in MoDCs pooled from 13 donors, and measured global gene expression by microarray (**Supplementary Fig. 2.1b**). We selected time points near the peaks of the kinetic curves of induced genes (5 hrs for LPS stimulation, and 10 hrs for influenza ΔNS1 infection).

We estimated the robustness of the technical components of the assay by performing separate cellular isolations from the same PBMC source, and by performing separate stimulations from the same MoDC source. The gene expression of LPS-stimulated MoDCs derived from separate cellular isolations from the same PBMC source showed a correlation of $R^2 = 0.98$ (**Supplementary Fig. 2.1c**). The gene expression of LPS-stimulated MoDCs derived from the same MoDC source showed a correlation of $R^2 = 0.99$ (**Supplementary Fig. 2.1d**). These results suggested that the assay is technically robust.

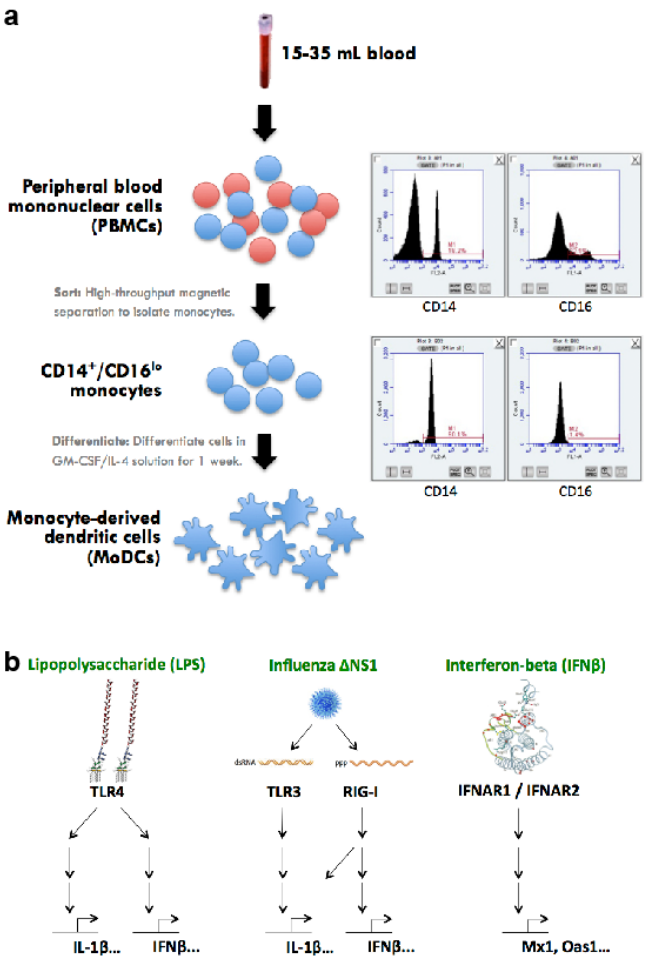


Figure 2.1. A robust system to quantitatively assess innate immune responses in humans.
(**a**) Peripheral blood from healthy donors was subjected to Ficoll centrifugation to isolate PBMCs, which were subjected to high-throughput magnetic sorting to isolate $CD14^{+}CD16^{lo}$ monocytes. Representative flow cytometry graphs of PBMCs and monocytes stained with anti-CD14 and anti-CD16 antibodies is shown. The monocytes were then differentiated into MoDCs for 1 week with GM-CSF and IL-4. (**b**) Purified LPS, influenza A ΔNS1, and recombinant IFNβ were used to stimulate the MoDCs. The stimuli engage the TLR4, RIG-I and TLR3, and IFNAR pathways, respectively.

We then devised a two-stage design to characterize innate immune variation between individuals. In stage 1, we collected genome-wide expression data for 30 donors and 9 serial replicates. From this data, we selected genes that showed evidence of reproducible expression variation and response variation. In stage 2, we collected gene expression data for about 400 selected genes in 133 donors, and performed an association study using genome-wide genotype data.

Comparison of LPS- and influenza-stimulated gene expression among 30 donors

First, we assessed the level of variation in gene expression of MoDCs from 30 healthy individuals (**Fig. 2.2a**) at baseline and in response to LPS and influenza ΔNS1. Following stimulation, we extracted RNA from the untreated and stimulated MoDCs, and then characterized genome-wide gene expression profiles using Affymetrix Human Gene 1.0 ST arrays. We normalized the expression data using RMA, and subtracted out the effect of sex and race using a linear mixed model. Using a pairwise t-test, we classified 837 and 1359 genes as significantly ($P < 0.05$) upregulated at $\log_2$(fold change) > 1.0 in response to LPS stimulation or influenza ΔNS1 infection, respectively; and 333 and 549 genes, respectively, at $\log_2$(fold change) > 1.5 (**Fig. 2.2b**). We found that the upregulated genes were enriched for Gene Ontology (GO) biological processes relevant to host defense such as "I-kappaB kinase/NF-kappaB cascade (GO:0007249)", "JAK-STAT cascade (GO:0007259)", and "cellular defense response (GO:0006968)". We found that the downregulated genes were enriched for metabolic and other non-specific pathways, which may reflect morphological or metabolic changes in the cell after stimulation (consequently, we gave less focus to downregulated genes in subsequent experiments).

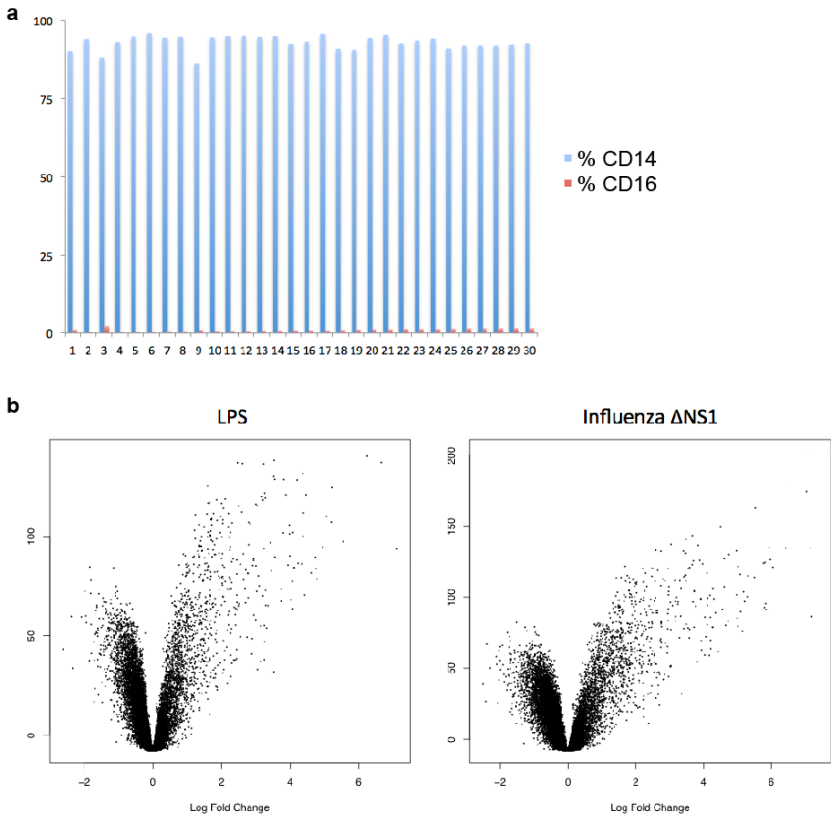


Figure 2.2. Stimulation of human MoDCs with LPS and influenza ΔNS1. (**a**) Flow cytometry analysis of monocytes isolated from 30 donors and stained with anti-CD14 and anti-CD16 antibodies. (**b**) Volcano plots showing differentially expressed genes after LPS stimulation (15 ng/mL) for 5 h, or influenza ΔNS1 infection for 10 h. The negative log transformed *P* values (y-axis) test the null hypothesis of no difference in expression levels between untreated and stimulated DCs and are plotted against the average log fold changes in expression (x-axis).

For each gene, we calculated the LPS and influenza ΔNS1 response values by calculating the difference between the log expression after stimulation and at baseline. We then calculated the variance of the (i) baseline expression values, (ii) LPS-induced expression values, (iii) influenza ΔNS1-induced expression values, (iv) LPS response values, and (v) influenza ΔNS1 response values. While some genes showed little inter-individual variation among the 30 individuals, other genes showed extensive variation (**Fig. 2.3a**).

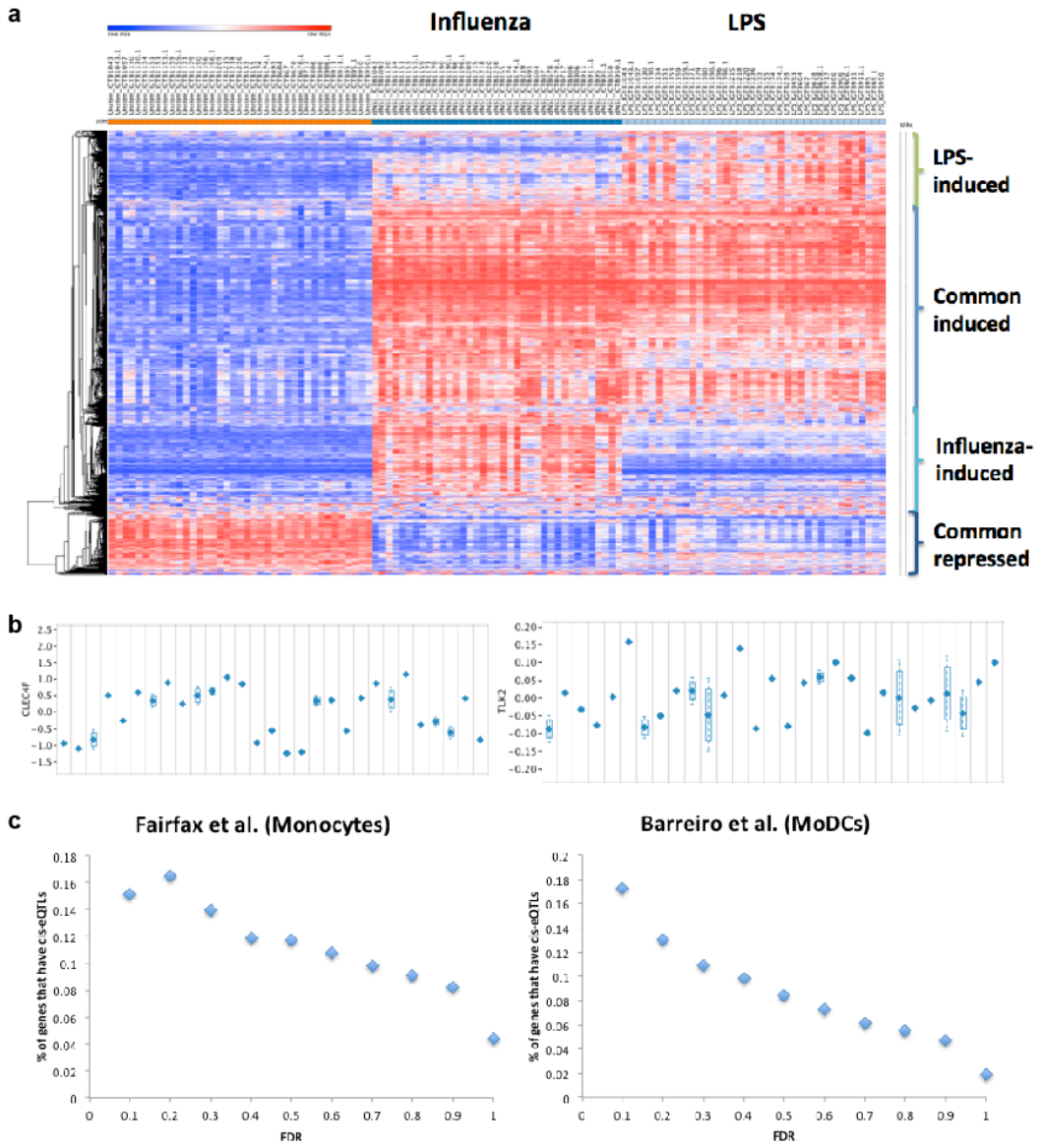


Figure 2.3. Inter- versus intra-individual variation in gene expression. (**a**) Heatmap of selected genes in MoDCs of 30 donors. Each column represents a donor sample at baseline, infected with influenza, or stimulated with LPS. Each row represents a gene. (**b**) Selected genes that showed variability in the LPS response. The graphs show variation in LPS responses of *CLEC4F* and *TLK2* among the individuals; standard deviations are marked for serial replicates. *CLEC4F* shows good serial reproducibility, while *TLK2* shows poor serial reproducibility. (**c**) FDR enriches for known cis-eQTLs. The x-axis represents FDR cutoffs for baseline expression data, and the y-axis represents percent of the genes that were found to have statistically significant cis-eQTLs in the baseline expression data in other studies (Barreiro et al. 2012; Fairfax et al. 2012).

Comparison of LPS- and influenza-stimulated gene expression within 9 serial replicates

The observed variation in gene expression may be due to underlying genetic variation, or to non-genetic factors such as technical noise (e.g. in the cellular isolation or stimulation) or biological noise (e.g. expression changes from molecules in the donor's blood). We hypothesized that culturing and differentiating the cells for seven days in a defined medium may reduce biological noise compared to lysing cells immediately after collection.

To directly assess the reproducibility of the expression and response values, we conducted a serial replication experiment. We collected fresh blood samples from 9 of the 30 donors six to nine months after the first collection. We repeated the cellular isolations, stimulations, and microarrays using the same assays and conditions as the initial experiment. We normalized the microarray data using RMA, and again subtracted out the effects of sex and race. We also corrected for the batch effect using surrogate variable analysis.

For each gene, we sought to determine whether the variation in expression and response within the serially replicated samples was less than the variation in expression and response among the different individuals. We calculated the false discovery rate (FDR) of inter- vs. intra-individual variation using a linear mixed model. We identified genes that demonstrated significant (FDR < 0.1) inter- vs. intra-individual variation at baseline, after LPS stimulation, after influenza ΔNS1infection, in response to LPS stimulation, and in response to influenza ΔNS1 infection (**Fig. 2.3b**). These results suggested that certain innate immune responses are highly reproducible and that there could be a substantial genetic component underlying them.

In order to further assess whether the genes with robust inter- vs intra-individual FDR scores may have an underlying genetic component, we determined whether these genes were enriched for known cis-eQTLs. We determined the number of genes at various FDR cutoffs that

were found to have cis-eQTLs in two other studies: one of monocytes (Fairfax et al. 2012) and one of MoDCs (Barreiro et al. 2012). We found that the genes with low inter- vs. intra-individual FDR scores in the baseline expression values are indeed enriched for cis-eQTLs found in other studies (**Fig. 2.3c**). These results further suggested that the genes with serially reproducible expression and response values (i.e. low FDR values) are likely to be associated with underlying genetic variants.

Because our cohort contained three different populations – Caucasian, East Asian, and African-American – we also assessed the robustness of variation among the three populations. We calculated the FDR of inter- vs. intra-population variation using a linear mixed model, and identified genes that demonstrated significant (FDR < 0.1) inter- vs. intra-population variation in response to LPS stimulation or in response to influenza ΔNS1 infection.

Selection of reporter genes

In order to further assess variation in innate immune responses in a larger cohort of individuals, we first created a reporter gene codeset (**Fig. 2.4a**). First, we selected all genes with a $\log_2$(fold change) greater than 0.5 or less than -1.5, and with an inter- vs. intra-individual FDR < 0.1 in (i) LPS-induced expression values, (ii) influenza ΔNS1-induced expression values, (iii) LPS response values, and (iv) influenza ΔNS1 response values. 97 of these genes were shared between conditions, while 125 were unique to one condition; in total we added 222 non-overlapping genes to the reporter gene codeset that showed evidence of serial reproducibility. We added 19 genes with known cis- or trans-eQTLs in monocytes or MoDCs as positive controls (Barreiro et al. 2012; Fairfax et al. 2012). To capture population variation, we added 21 genes

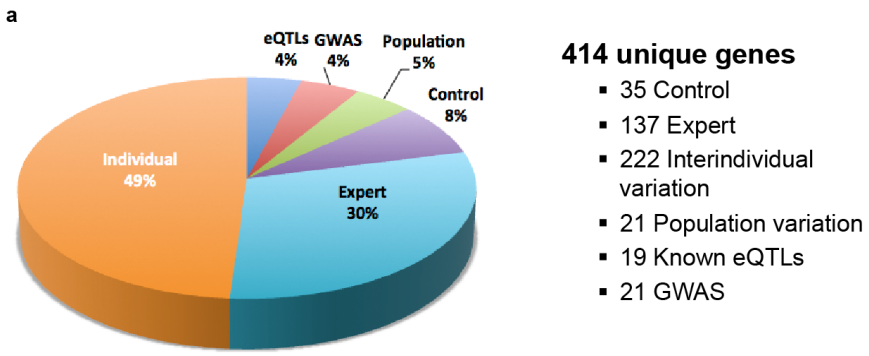


Figure 2.4. Selection of reporter genes for further interrogation. (**a**) Pie chart showing genes represented on Nanostring codeset. The codeset contains a total of 414 unique genes with 222 representing inter- vs. intra-individual variation genes; 21 representing inter- vs. intra-population variation genes; 137 curated (expert) genes including known pathway components and key cytokines; 19 known eQTLs and 21 known GWAS hits from infectious and autoimmune disease studies; and 35 control genes including low variance genes, sex-specific genes, and non-expressed genes.

with a $\log_2$(fold change) greater than 0.5 or less than -1.5, and with an inter- vs. intra-population FDR < 0.1 in LPS response values or influenza response values.

In addition to this set of genes, we added 137 genes that are known players in the relevant pathways. These included the receptors *TLR4*, *TLR3*, *RIG-I*, *IFNAR1*, and *IFNAR2*; adaptors including *MYD88*, *TRIF*, *STING*, *MAVS*, *JAK1*, *STAT1*, and *TYK2*; and key regulated outputs including *IFNB1*, *IL28B*, *TNF*, *IL6*, *IL10,* and *IFITM3*. We also added genes such as *IL12B*, *ACP5*, and *STAT3* that are expressed in MoDCs and cause Mendelian infectious or autoimmune diseases when mutated.

As controls, we added 15 genes (e.g. *GAPDH*) with low variance across the microarray data, and 8 genes that were regulated but had response values with low variance. To control for contaminated cell types, we added genes that are not expressed in MoDCs including *CD3*, *CD19*, *CD56*, and *CD235a*. To detect possible sample switching, we added sex-specific genes from the Y-chromosome including *DDX3Y*, *EIF1AY,* and *ZFY*. In total, we selected 414 unique genes for

our Nanostring codeset (**Fig. 2.4a**). We chose probes that excluded common SNPs (minor allele frequency > 5%) except for 11 genes for which this was not possible.

Comparison of LPS-, influenza-, and IFNβ-stimulated gene expression in 133 donors

We then collected blood samples and isolated PBMCs from 643 healthy donors with 56 additional serial replicates (i.e. separate blood draws > 1 month later). We isolated genomic DNA from each sample and genotyped the DNA on the Illumina HumanOmni1-Quad BeadChip and Illumina Infinium HumanExome BeadChip to type about 1.1 million SNPs and 240,000 exonic variants, respectively.

We prepared and cultured MoDCs from 133 of the samples so far. In addition to a baseline sample, we stimulated the cells with LPS, influenza ΔNS1, or recombinant interferon-beta (IFNβ), for up to 4 conditions per donor depending on available cell number. We compared gene expression and response values for these donors (**Fig. 2.5a**). Preliminary analysis suggests that there are significant cis-eQTLs in the baseline expression, stimulated expression, and response values for all stimuli (**Fig. 2.5b**). Several of these cis-eQTLs, including *IRF5*, *IFITM3*, and *XBP1*, have been associated with infectious or autoimmune phenotypes, including systemic lupus erythematosus, rheumatoid arthritis, influenza morbidity, and inflammatory bowel disease (Graham et al. 2006; Kaser et al. 2008; Stahl et al. 2010; Everitt et al. 2012).

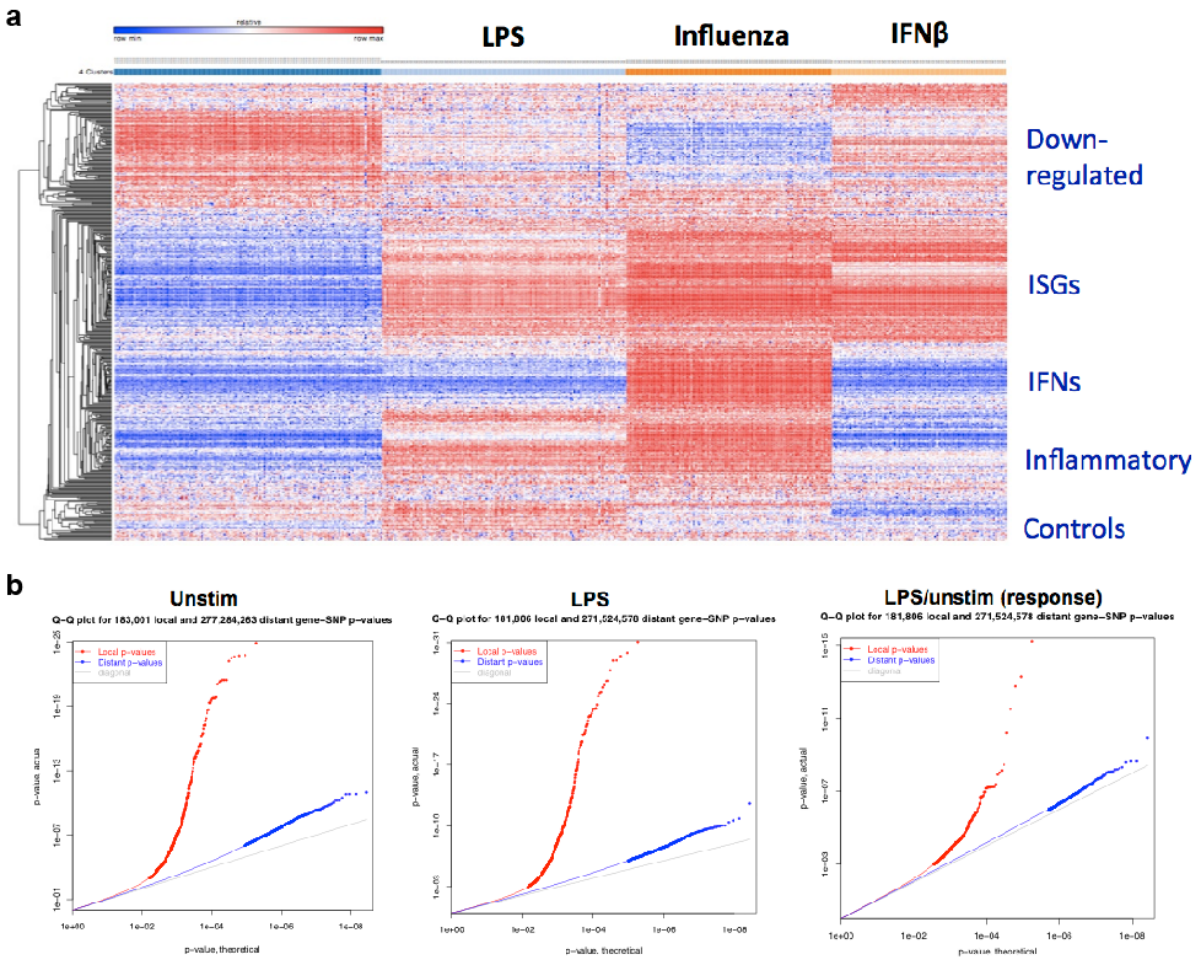


Figure 2.5. Variation in gene expression of MoDCs from 133 donors. (**a**) Heatmap of 414 reporter genes in MoDCs of 133 donors. Each column represents a donor sample at baseline or stimulated with LPS, influenza, or IFNβ. Each row represents the expression of a gene. (**b**) Preliminary eQTL analysis of selected gene expression values from 133 donors (unstim, LPS, and LPS response). Q-Q plots show deviation of actual *P*-values (y-axis) from expected *P*-values (x-axis) for cis-eQTLs (red). The significance of the trans-eQTL signals (blue) are less clear.

DISCUSSION

Variation in gene expression and to a more limited extent, variation in the change in expression in response to a stimulus has been observed in humans for a number of cell types (Storey et al. 2007). When expression and genotype data are collected for a large number of individuals, studies have been able to associate individual genetic variants to the observed expression (Stranger et al. 2007) and response variation (Smirnov et al. 2009; Romanoski et al. 2010; Barreiro et al. 2012) with statistical robustness. Baseline studies comparing eQTLs in monocytes and B cells identified numerous genes with genetic variants associated with the expression of the gene in a cell type specific fashion (Fairfax et al. 2012).

Our study differs from previous studies in that (i) multiple stimuli affecting distinct components of the TLR signaling pathway are used and (ii) serial replicates are collected for a limited number of individuals to identify the effects of short term environmental differences, stimulation, and genetic variation on expression variation. This work has implications for decomposing the proportion of variance explained by environmental and genetic factors and understanding the different genetic and environmental effects on TLR response components.

We show that unlike previous work, the estimates for inter-individual gene expression is often inflated when only technical replicates are considered. However, both the baseline and response expression of a large number of genes are still remarkably reproducible between serial replicates collected months apart. We are able to identify key components in both the bacterial and viral response pathways that are both consistent within an individual and variable between individuals. We show that these results are highly consistent with previous results based on mapping studies. We further validated our results by performing a preliminary mapping study and identified a number of significant cis-eQTLs.

With further analysis, by combining detailed characterization of innate immune phenotypes with genotyping data, we can determine how different genotypes influence innate immune responses in humans. Using this collection of variants as a resource, we could ask whether these variants play a role in various infectious and autoimmune phenotypes in humans *in vivo*. Our results may also point to a molecular mechanism underlying GWAS hits of unknown function.

MATERIALS AND METHODS

Reagents. LPS (ultrapure lipopolysaccharide from *E. coli* K12) was obtained from Invivogen. Influenza ΔNS1 was prepared as described (Shapira et al. 2009). Recombinant human IFNβ was obtained from PBL Interferon Source.

Study subjects. Donors were recruited from the Boston community and gave written informed consent for the studies. Individuals were excluded if they had a history of inflammatory disease, autoimmune disease, chronic metabolic disorders, or chronic infectious disorders. For the microarray study, 30 healthy human donors were recruited. Donors were between 19 and 49 years of age; 15 were male and 15 were female; 18 were Caucasian, 6 were East Asian, and 6 were African-American. For the Nanostring study, 643 healthy human donors were recruited. Donors were between 18 and 56 years of age (avg. 29.9); 285 were male and 358 were female; 363 were Caucasian, 148 were East Asian, 164 were African-American, and 24 were multi-racial.

Preparation and stimulation of primary human monocyte-derived dendritic cells (MoDCs). 35-50 mL of peripheral blood from fasting subjects was collected between 7:30 am and 8:30 am. The blood was drawn into sodium heparin tubes and peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque (GE Healthcare Life Sciences) centrifugation. PBMCs were frozen in liquid N_2 in 90% FBS (Sigma) and 10% DMSO (Sigma). Monocytes were isolated from PBMCs by negative selection using the Dynabeads Untouched Human Monocytes kit (Life Technologies) modified to increase throughput and optimize recovery and purity of $CD14^{+}CD16^{lo}$ monocytes: the FcR Blocking Reagent was replaced with Miltenyi FcR Blocking Reagent (Miltenyi); per mL of Antibody Mix, an additional 333 ug biotinylated anti-CD16 (3G8),

167 ug biotinylated anti-CD3 (SK7), and 167 ug biotinylated anti-CD19 (HIB19) antibodies (Biolegend) were added; the antibody labeling was performed in 96-well plates; and Miltenyi MS Columns or Multi-96 Columns (Miltenyi) were used to separate the magnetically-labeled cells from the unlabeled cells in an OctoMACS Separator or MultiMACS M96 Separator (Miltenyi). The number of PBMCs and monocytes was estimated using CellTiter-Glo Luminescent Cell Viability Assay (Promega). A subset of the isolated monocytes was stained with PE-labeled anti-CD14 (M5E2; BD Biosciences) and FITC-labeled anti-CD16 (3G8; Biolegend), and subjected to flow cytometry analysis using an Accuri C6 Flow Cytometer (BD Biosciences). The remaining monocytes were cultured for seven days in RPMI (Life Technologies) supplemented with 10% FBS (Sigma), 100 ng/mL GM-CSF (R&D), and 40 ng/mL IL-4 (R&D) to differentiate them into monocyte-derived dendritic cells (MoDCs). $5x10^4$ MoDCs were seeded in each well of a 96-well plate, and stimulated with 15 ng/mL LPS for 5 h, influenza ΔNS1 virus for 10 h, 100 U/mL recombinant human IFNβ for 6.5 h, or left unstimulated. Cells were then lysed in RLT buffer (Qiagen) supplemented with β-mercaptoethanol (Sigma).

Microarray gene expression profiling. Total RNA was isolated using the RNeasy Mini kit (Qiagen). RNA quantification and quality were assessed using Agilent 2100 Bioanalyzer (Agilent Technologies). The cDNA synthesis, labeling, and subsequent hybridization to the microarrays were performed by the company Expression Analysis (Durham, NC). The Affymetrix Human Gene 1.0 ST arrays were used to obtain genome-wide gene expression profiles.

Normalization of microarray data and adjustment for confounding effects. Outlier detection was first performed. The dataset was then normalized using quantile normalization as part of the RMA pipeline. To eliminate possible confounding affects, in particular the batch effect from running our array over two plates, surrogate variable analysis was used to identify a number of highly significant surrogate variables (SVs) identified through a permutation test and these were included in our follow up analysis.

Mixed model estimate for proportion of variance explained by stimulation, environment and genetic variants. We used a linear mixed model to model the variance components for the expression of an individual gene:

$y_i = a \times LPS + b \times dNS1 + c \times Gender + d \times Age + \alpha \times Interindividual + \beta \times Intraindividual + s_1 \times SV_1 + s_2 \times SV_2 + \ldots s_n \times SV_n + \varepsilon$

Mixed model estimates for inter-individual and inter-population variable genes. To test for interindividual variability, we tested the above model against the null where inter-individual variability is estimated to be 0 ($\alpha = 0$).

Reporter gene selection. Normalization controls were selected by choosing genes with low variance in expression in the microarray data. Additional low variance controls were selected by choosing genes with low variance in response to LPS or influenza ΔNS1 stimulation. Positive controls that represented genes with known cis-eQTLs were selected from baseline monocyte cis-eQTLs in Fairfax *et al*. (Fairfax et al. 2012) and from baseline MoDC cis-eQTLs in Barreiro *et al*. (Barreiro et al. 2012). Known pathway components were selected from KEGG, Ingenuity

Pathway Analysis, and references including (Akira and Takeda 2004; Barbalat et al. 2011). We added genes that are not expressed in MoDCs including *CD3*, *CD19*, *CD56*, and *CD235a* to control for contaminated cell types. We added sex-specific genes from the Y-chromosome including *DDX3Y*, *EIF1AY,* and *ZFY* to detect possible sample switching. To capture inter-individual variation, we added genes with a $\log_2$(fold change) > 0.5 or $\log_2$(fold change) < -1.5 (biased because downregulated genes are of less clear function by GO categories), and with an inter- vs. intra-individual FDR < 0.1 in (i) LPS-induced expression values, (ii) influenza ΔNS1-induced expression values, (iii) LPS response values, and (iv) influenza ΔNS1 response values. To capture variation in population, we added genes with a $\log_2$(fold change) greater than 0.5 or less than -1.5, and with an inter- vs. intra-population FDR < 0.1 in LPS response values and influenza ΔNS1 response values. We did not include genes that had a maximum value of $\log_2$ > 13 across the microarray dataset. We chose probes that excluded common SNPs (minor allele frequency > 5%) except for 11 genes for which this was not possible.

Nanostring. The Nanostring nCounter system (Nanostring) was used to digitally count transcripts in a multiplex reaction as previously described (Geiss et al. 2008). Lysates in RLT buffer were hybridized for 12-24 hours with custom nCounter Gene Expression CodeSets. Quantification of hybridized RNA was performed using the nCounter Analysis System. To normalize the nCounter data, we first normalized each sample using the internal positive spike-in controls, and then normalized the data based on the average expression level of the low variance control genes. Surrogate variable analysis was used to eliminate possible confounding effects.

CONCLUSION

Main findings

Aim #1: Identify novel ISD pathway components. By precipitating DNA from a cytoplasmic lysate, we identified 184 candidate cytoplasmic DNA-interacting proteins. 10% of these proteins are known to be involved in innate immune pathways already, and the others may be useful as a resource for study of innate immune sensing as well as host-viral interaction. We silenced 809 candidate genes, measured the response to dsDNA, and connected resulting hits with the known signaling network. We identified many novel genes that decreased or increased the ISD response after knockdown; these may be useful as a resource for identification of new pathway members. We validated ABCF1, and demonstrated that it is a critical protein that associates with dsDNA and the known DNA-sensing components, HMGB2 and IFI16. We also found that CDC37 regulates stability of the signaling molecule, TBK1, and that chemical inhibition of CDC37 and several other pathway regulators (HSP90, TBK1, PTPN1, PPP6C) potently modulates the innate immune response to DNA and to retroviral infection in *Trex1*-/- MEFs as well as in AGS patient cells. These molecules may be therapeutic leads in *TREX1*-mediated disorders such as AGS.

Aim #2: Characterize variation in innate immune responses in primary dendritic cells in humans, and determine whether the functional variation is caused by genetic variation. We developed a high-throughput system to quantitatively measure innate immune responses in primary human MoDCs. We characterized variation in responses between 30 healthy donors, and described which variants are reproducible using a serial replicate design. We then developed a reporter gene codeset, characterized variation in responses between 133 healthy donors, and performed an initial association study with corresponding genotype data. We found

cis-eQTLs in genes in response to the stimuli, several of which are known to be associated with infectious and autoimmune clinical phenotypes.

The strengths and weaknesses of approaches taken

(i) Proteomics.

Strengths: The proteomics approach that we used to identify candidate cytoplasmic DNA-interacting proteins and candidate STING-interacting proteins was both unbiased and quantitative. The SILAC approach allowed us to find many more proteins than we could visualize by PAGE (the problem with cutting out bands after PAGE is low sensitivity; additionally, protein bands can be obscured by non-specific background bands).

Weaknesses: Proteomics approaches have some false positive rate due to non-physiological interactions that occur *in vitro* after lysing the whole cell and concentrating the lysate in buffer. This was perhaps more markedly seen with our STING pulldowns, as most of the hits did not have a functional effect in our siRNA screen. We also had some false negative rate. For example, we did not identity the interaction of TBK1 with STING. Some of the false negatives likely occurred because of transient interactions that are difficult to capture, in addition to us not capturing the correct time point after stimulation.

Alternative approaches: In the future, crosslinking could be used to stabilize transient interactions. More time points can be taken after stimulation. Different amounts of ligands can be used. More controls can be pulled down for comparison.

(ii) RNAi screening.

Strengths: The RNAi screening approach that we used was a high-throughput method that allowed us to test many hypotheses in parallel.

Weaknesses: RNAi has a clear false positive rate, though it is hard to quantitate. Originally we wanted to estimate the false positive rate by performing a large-scale rescue experiment of our hits, but this was more technically difficult than we anticipated. RNAi also has some false negative rate, which depends on the screening library. In our screen, we captured expected components including *Tbk1*, *Cxcl10*, *Jak1*, *Stat1*, and *Irf9*. We also captured *Irf3* but only because we used a different siRNA than was in the Dharmacon SMARTpool siRNA library. The SMARTpool siIRF3 was toxic; this was off-target toxicity, which is one source of false negatives. We captured *Sting* but only because we added the siRNA into the screen. There was no siRNA targeting *Sting* in the Dharmacon library we used; this is another source of false negatives. We missed *Ifi16* since the siRNA did not have a phenotype in our screen. Later tests using a panel of different siRNAs demonstrated the partial phenotype shown by the other groups.

Alternative approaches: Two or more different siRNA libraries can be used to ameliorate the false negative and false positive rates seen here, though there is a cost in terms of both money and effort to do this. Alternative screening tools such as haplotype screening (Carette et al. 2011), ENU mutagenesis (Tabeta et al. 2006), overexpression screening (Kawai et al. 2005), and homology-based approaches (Meylan et al. 2005; Seth et al. 2005) may have a smaller false positive rate.

(iii) eQTL analysis.

Strengths: Because genetic variants must be causal, genetic association studies have been a powerful method to map phenotypes. Nanostring is technically simple and highly reproducible.

Weaknesses: We were limited by our readouts. The microarrays were only performed on 30 donors, and do not capture all pertinent information about the expressed genes, such as isoforms. Nanostring only measured gene expression of a set of hundreds of genes. Genotyping only measures a fraction of genetic variants, albeit the most common ones.

Alternative approaches: RNA sequencing and DNA sequencing captures all the information that we lost, albeit at higher cost.

Future studies

ISD pathway. Our RNAi screening assay could be scaled up to test nearly the whole genome (i.e. the entire siRNA library). This is very practical in terms of cost and effort at this point, since the assay has already been set up. This would be a more unbiased approach to identifying novel pathway members. The downside is that, because of the false positive rate, it may be difficult to determine which hits are on-target. Thus, parallel to this approach, more datasets could be generated like the SILAC datasets described above. The STING pulldown could be improved as described above. Additionaly, proteins that are post-translationally modified (e.g. phosphorylated or ubiquitinated) could be globally measured and used to inform the functional data (Chevrier et al. 2011).

There are a number of questions that remain about the mechanism by which ABCF1 functions in the ISD pathway. Is the interaction with DNA direct or indirect? Does ABCF1 interact with HMGB2 and IFI16 in a single complex or in multiple complexes? Does ABCF1

interact with downstream signaling molecules such as STING? If not, what is the role of ABCF1 in the DNA sensing complex? We had attempted to answer the first three questions, but were not able to generate convincing data to answer them at this point, and further work is needed.

In regards to the small molecule inhibitors, we are planning to test them *in vivo* in the *Trex1*-/- animal model. *Trex1*-/- mice develop severe autoimmune disease notably characterized by an inflammatory myocarditis. We will inject each of the compounds into *Trex1*-/- mice intraperitoneally, with vehicle as control. Our readouts will be *Ifnb1* mRNA levels in the heart tissue of mice, as well as survival. We expect to see decreased *Ifnb1* expression in mice treated with small molecules, as well as increased survival times relative to control mice.

eQTL project. Because we did not use all the RNA from the experiment, the remaining RNA could be sequenced in the future. We could either wait until costs drop or we could use a simplified protocol to sequence only the 3' ends. Because the design of the experiment and the assays are set up, more stimuli and more cell types could be used to generate similar types of data. Theoretically, eQTL maps should be made for every cell type x every stimulated condition that is relevant *in vivo*.

Similar to GWAS designs, a case-cohort design could be used to ask whether there are differences in innate immune responses in the cells of healthy individuals vs. individuals with infectious or autoimmune diseases. Given the data described in the introduction, one can hypothesize that individuals with certain types of diseases will have altered innate immune responses, which could be captured by a case-control design. The upside of a functional approach as described here is that it could capture any genetic variation – common, rare, or synergistic – in the network (without needing prior knowledge of all its components). This may be an advantage over genetic studies which, for example, have difficulties capturing gene x gene

interactions. The downside that the cells may be altered by the disease process (e.g. a pathogen) or by drugs that the patient is taking, and this may confound the results.

Concluding remarks

Overall, we demonstrated how high-throughput genomic approaches can be used to identify genes and genetic variants that are associated with human disease. Given the abundance of genetic data that is currently being collected to identify variants that associate with clinical phenotypes, there is an existing need for approaches like these to understand the molecular phenotypes of the disease-associating variants. Deciphering the molecular pathways and components that are altered in these diseases will hopefully aid in the development of therapeutics to target the pathways and components in patients.

REFERENCES

Ablasser A, Bauernfeind F, Hartmann G, Latz E, Fitzgerald KA, Hornung V. 2009. RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nat Immunol* **10**(10): 1065-1072.

Akira S, Takeda K. 2004. Toll-like receptor signalling. *Nat Rev Immunol* **4**(7): 499-511.

Akira S, Uematsu S, Takeuchi O. 2006. Pathogen recognition and innate immunity. *Cell* **124**(4): 783-801.

Alexopoulou L, Holt AC, Medzhitov R, Flavell RA. 2001. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* **413**(6857): 732-738.

Altfeld M, Fadda L, Frleta D, Bhardwaj N. 2011. DCs and NK cells: critical effectors in the immune response to HIV-1. *Nat Rev Immunol* **11**(3): 176-186.

Amit I, Garber M, Chevrier N, Leite AP, Donner Y, Eisenhaure T, Guttman M, Grenier JK, Li W, Zuk O et al. 2009. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* **326**(5950): 257-263.

Audry M, Ciancanelli M, Yang K, Cobat A, Chang HH, Sancho-Shimizu V, Lorenzo L, Niehues T, Reichenbach J, Li XX et al. 2011. NEMO is a key component of NF-kappaB- and IRF-3-dependent TLR3-mediated immunity to herpes simplex virus. *The Journal of allergy and clinical immunology* **128**(3): 610-617 e611-614.

Barbalat R, Ewald SE, Mouchess ML, Barton GM. 2011. Nucleic acid recognition by the innate immune system. *Annual review of immunology* **29**: 185-214.

Barreiro LB, Ben-Ali M, Quach H, Laval G, Patin E, Pickrell JK, Bouchier C, Tichit M, Neyrolles O, Gicquel B et al. 2009. Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS genetics* **5**(7): e1000562.

Barreiro LB, Tailleux L, Pai AA, Gicquel B, Marioni JC, Gilad Y. 2012. Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection. *Proceedings of the National Academy of Sciences of the United States of America* **109**(4): 1204-1209.

Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C et al. 2009. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* **41**(6): 703-707.

Barton GM, Kagan JC, Medzhitov R. 2006. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nat Immunol* **7**(1): 49-56.

Beck-Engeser GB, Eilat D, Wabl M. 2011. An autoimmune disease prevented by anti-retroviral drugs. *Retrovirology* **8**: 91.

Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, Pascual V. 2003. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *The Journal of experimental medicine* **197**(6): 711-723.

Bogunovic D, Byun M, Durfee LA, Abhyankar A, Sanal O, Mansouri D, Salem S, Radovanovic I, Grant AV, Adimi P et al. 2012. Mycobacterial disease and impaired IFN-gamma immunity in humans with inherited ISG15 deficiency. *Science* **337**(6102): 1684-1688.

Bonnard M, Mirtsos C, Suzuki S, Graham K, Huang J, Ng M, Itie A, Wakeham A, Shahinian A, Henzel WJ et al. 2000. Deficiency of T2K leads to apoptotic liver degeneration and impaired NF-kappaB-dependent gene transcription. *The EMBO journal* **19**(18): 4976-4985.

Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C, Hurley P, Chien M, Chai S, Hitotsumatsu O et al. 2004. The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat Immunol* **5**(10): 1052-1060.

Bouwmeester T, Bauch A, Ruffner H, Angrand PO, Bergamini G, Croughton K, Cruciat C, Eberhard D, Gagneur J, Ghidelli S et al. 2004. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* **6**(2): 97-105.

Burckstummer T, Baumann C, Bluml S, Dixit E, Durnberger G, Jahn H, Planyavsky M, Bilban M, Colinge J, Bennett KL et al. 2009. An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nat Immunol* **10**(3): 266-272.

Caetano BC, Carmo BB, Melo MB, Cerny A, dos Santos SL, Bartholomeu DC, Golenbock DT, Gazzinelli RT. 2011. Requirement of UNC93B1 reveals a critical role for TLR7 in host resistance to primary infection with Trypanosoma cruzi. *Journal of immunology* **187**(4): 1903-1911.

Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, Kuehne AI, Kranzusch PJ, Griffin AM, Ruthel G et al. 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**(7364): 340-343.

Casanova JL, Abel L. 2004. The human model: a genetic dissection of immunity to infection in natural conditions. *Nat Rev Immunol* **4**(1): 55-66.

Casanova JL, Abel L, Quintana-Murci L. 2011. Human TLRs and IL-1Rs in host defense: natural insights from evolutionary, epidemiological, and clinical genetics. *Annual review of immunology* **29**: 447-491.

Casanova JL, Holland SM, Notarangelo LD. 2012. Inborn errors of human JAKs and STATs. *Immunity* **36**(4): 515-528.

Casrouge A, Zhang SY, Eidenschenk C, Jouanguy E, Puel A, Yang K, Alcais A, Picard C, Mahfoufi N, Nicolas N et al. 2006. Herpes simplex virus encephalitis in human UNC-93B deficiency. *Science* **314**(5797): 308-312.

Chevrier N, Mertins P, Artyomov MN, Shalek AK, Iannacone M, Ciaccio MF, Gat-Viks I, Tonti E, DeGrace MM, Clauser KR et al. 2011. Systematic discovery of TLR signaling components delineates viral-sensing circuits. *Cell* **147**(4): 853-867.

Chiu YH, Macmillan JB, Chen ZJ. 2009. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* **138**(3): 576-591.

Cho JH. 2008. The genetics and immunopathogenesis of inflammatory bowel disease. *Nat Rev Immunol* **8**(6): 458-466.

Chowdhury D, Lieberman J. 2008. Death by a thousand cuts: granzyme pathways of programmed cell death. *Annual review of immunology* **26**: 389-420.

Coccia EM, Severa M, Giacomini E, Monneron D, Remoli ME, Julkunen I, Cella M, Lande R, Uze G. 2004. Viral infection and Toll-like receptor agonists induce a differential expression of type I and lambda interferons in human plasmacytoid and monocyte-derived dendritic cells. *European journal of immunology* **34**(3): 796-805.

Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. 2009. Mapping complex disease traits with global gene expression. *Nature reviews Genetics* **10**(3): 184-194.

Corson TW, Crews CM. 2007. Molecular understanding and modern application of traditional medicines: triumphs and trials. *Cell* **130**(5): 769-774.

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* **10**(4): 1794-1805.

Cristea IM, Moorman NJ, Terhune SS, Cuevas CD, O'Keefe ES, Rout MP, Chait BT, Shenk T. 2010. Human cytomegalovirus pUL83 stimulates activity of the viral immediate-early promoter through its interaction with the cellular IFI16 protein. *Journal of virology* **84**(15): 7803-7814.

Crow YJ. 2011. Type I interferonopathies: a novel set of inborn errors of immunity. *Annals of the New York Academy of Sciences* **1238**: 91-98.

Crow YJ, Hayward BE, Parmar R, Robins P, Leitch A, Ali M, Black DN, van Bokhoven H, Brunner HG, Hamel BC et al. 2006. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. *Nat Genet* **38**(8): 917-920.

DeFilippis VR, Alvarado D, Sali T, Rothenburg S, Fruh K. 2010. Human cytomegalovirus induces the interferon response via the DNA sensor ZBP1. *J Virol* **84**(1): 585-598.

Dermitzakis ET. 2012. Cellular genomics for complex traits. *Nature reviews Genetics* **13**(3): 215-220.

Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M et al. 2009. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325**(5945): 1246-1250.

Dixit E, Boulant S, Zhang Y, Lee AS, Odendall C, Shum B, Hacohen N, Chen ZJ, Whelan SP, Fransen M et al. 2010. Peroxisomes are signaling platforms for antiviral innate immunity. *Cell* **141**(4): 668-681.

Dupuis S, Dargemont C, Fieschi C, Thomassin N, Rosenzweig S, Harris J, Holland SM, Schreiber RD, Casanova JL. 2001. Impairment of mycobacterial but not viral immunity by a germline human STAT1 mutation. *Science* **293**(5528): 300-303.

Dupuis S, Jouanguy E, Al-Hajjar S, Fieschi C, Al-Mohsen IZ, Al-Jumaah S, Yang K, Chapgier A, Eidenschenk C, Eid P et al. 2003. Impaired response to interferon-alpha/beta and lethal viral disease in human STAT1 deficiency. *Nat Genet* **33**(3): 388-391.

Everitt AR, Clare S, Pertel T, John SP, Wash RS, Smith SE, Chin CR, Feeley EM, Sims JS, Adams DJ et al. 2012. IFITM3 restricts the morbidity and mortality associated with influenza. *Nature* **484**(7395): 519-523.

Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO, Knight JC. 2012. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet* **44**(5): 502-510.

Fernandes-Alnemri T, Yu JW, Datta P, Wu J, Alnemri ES. 2009. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458**(7237): 509-513.

Fitzgerald KA, McWhirter SM, Faia KL, Rowe DC, Latz E, Golenbock DT, Coyle AJ, Liao SM, Maniatis T. 2003. IKKepsilon and TBK1 are essential components of the IRF3 signaling pathway. *Nat Immunol* **4**(5): 491-496.

Fornarino S, Laval G, Barreiro LB, Manry J, Vasseur E, Quintana-Murci L. 2011. Evolution of the TIR domain-containing adaptors in humans: swinging between constraint and adaptation. *Molecular biology and evolution* **28**(11): 3087-3097.

Fu XY, Schindler C, Improta T, Aebersold R, Darnell JE, Jr. 1992. The proteins of ISGF-3, the interferon alpha-induced transcriptional activator, define a gene family involved in signal

transduction. *Proceedings of the National Academy of Sciences of the United States of America* **89**(16): 7840-7843.

Gack MU, Albrecht RA, Urano T, Inn KS, Huang IC, Carnero E, Farzan M, Inoue S, Jung JU, Garcia-Sastre A. 2009. Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. *Cell Host Microbe* **5**(5): 439-449.

Gack MU, Shin YC, Joo CH, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S et al. 2007. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**(7138): 916-920.

Gall A, Treuting P, Elkon KB, Loo YM, Gale M, Jr., Barber GN, Stetson DB. 2012. Autoimmunity initiates in nonhematopoietic cells and progresses via lymphocytes in an interferon-dependent autoimmune disease. *Immunity* **36**(1): 120-131.

Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ et al. 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **461**(7262): 399-401.

Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T et al. 2008. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* **26**(3): 317-325.

Glocker EO, Frede N, Perro M, Sebire N, Elawad M, Shah N, Grimbacher B. 2010. Infant colitis--it's in the genes. *Lancet* **376**(9748): 1272.

Goldbach-Mansky R, Wilson M, Fleischmann R, Olsen N, Silverfield J, Kempf P, Kivitz A, Sherrer Y, Pucino F, Csako G et al. 2009. Comparison of Tripterygium wilfordii Hook F versus sulfasalazine in the treatment of rheumatoid arthritis: a randomized trial. *Ann Intern Med* **151**(4): 229-240, W249-251.

Gowen BB, Hoopes JD, Wong MH, Jung KH, Isakson KC, Alexopoulou L, Flavell RA, Sidwell RW. 2006. TLR3 deletion limits mortality and disease severity due to Phlebovirus infection. *Journal of immunology* **177**(9): 6301-6307.

Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C et al. 2008. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet* **40**(9): 1059-1061.

Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, Ortmann WA, Koeuth T, Gonzalez Escribano MF, Argentine et al. 2006. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* **38**(5): 550-555.

Graham RR, Kyogoku C, Sigurdsson S, Vlasova IA, Davies LR, Baechler EC, Plenge RM, Koeuth T, Ortmann WA, Hom G et al. 2007. Three functional variants of IFN regulatory

factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proceedings of the National Academy of Sciences of the United States of America* **104**(16): 6758-6763.

Gray PJ, Jr., Prince T, Cheng J, Stevenson MA, Calderwood SK. 2008. Targeting the oncogene and kinome chaperone CDC37. *Nat Rev Cancer* **8**(7): 491-495.

Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, Seldin MF, Criswell LA, Plenge RM, Holers VM et al. 2009. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat Genet* **41**(7): 820-823.

Guo Y, Audry M, Ciancanelli M, Alsina L, Azevedo J, Herman M, Anguiano E, Sancho-Shimizu V, Lorenzo L, Pauwels E et al. 2011. Herpes simplex virus encephalitis in a patient with complete TLR3 deficiency: TLR3 is otherwise redundant in protective immunity. *The Journal of experimental medicine* **208**(10): 2083-2098.

Hedl M, Abraham C. 2012. IRF5 risk polymorphisms contribute to interindividual variance in pattern recognition receptor-mediated cytokine secretion in human monocyte-derived cells. *Journal of immunology* **188**(11): 5348-5356.

Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, Matsumoto M, Hoshino K, Wagner H, Takeda K et al. 2000. A Toll-like receptor recognizes bacterial DNA. *Nature* **408**(6813): 740-745.

Herman M, Ciancanelli M, Ou YH, Lorenzo L, Klaudel-Dreszler M, Pauwels E, Sancho-Shimizu V, Perez de Diego R, Abhyankar A, Israelsson E et al. 2012. Heterozygous TBK1 mutations impair TLR3 immunity and underlie herpes simplex encephalitis of childhood. *The Journal of experimental medicine* **209**(9): 1567-1582.

Hoebe K, Du X, Georgel P, Janssen E, Tabeta K, Kim SO, Goode J, Lin P, Mann N, Mudd S et al. 2003. Identification of Lps2 as a key transducer of MyD88-independent TIR signalling. *Nature* **424**(6950): 743-748.

Hornung V, Ablasser A, Charrel-Dennis M, Bauernfeind F, Horvath G, Caffrey DR, Latz E, Fitzgerald KA. 2009. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458**(7237): 514-518.

Hornung V, Ellegast J, Kim S, Brzozka K, Jung A, Kato H, Poeck H, Akira S, Conzelmann KK, Schlee M et al. 2006. 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**(5801): 994-997.

Hyoty H, Taylor KW. 2002. The role of viruses in human diabetes. *Diabetologia* **45**(10): 1353-1361.

Isaacs A, Cox RA, Rotem Z. 1963. Foreign nucleic acids as the stimulus to make interferon. *Lancet* **2**(7299): 113-116.

Ishii KJ, Coban C, Kato H, Takahashi K, Torii Y, Takeshita F, Ludwig H, Sutter G, Suzuki K, Hemmi H et al. 2006. A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nat Immunol* **7**(1): 40-48.

Ishii KJ, Kawagoe T, Koyama S, Matsui K, Kumar H, Kawai T, Uematsu S, Takeuchi O, Takeshita F, Coban C et al. 2008. TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. *Nature* **451**(7179): 725-729.

Ishikawa H, Barber GN. 2008. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455**(7213): 674-678.

Ishikawa H, Ma Z, Barber GN. 2009. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461**(7265): 788-792.

Jacob CO, Zhu J, Armstrong DL, Yan M, Han J, Zhou XJ, Thomas JA, Reiff A, Myones BL, Ojwang JO et al. 2009. Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus. *Proceedings of the National Academy of Sciences of the United States of America* **106**(15): 6256-6261.

Janeway CA, Jr. 1989. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harbor symposia on quantitative biology* **54 Pt 1**: 1-13.

Jensen KE, Neal AL, Owens RE, Warren J. 1963. Interferon Responses of Chick Embryo Fibroblasts to Nucleic Acids and Related Compounds. *Nature* **200**: 433-434.

Jin L, Xu LG, Yang IV, Davidson EJ, Schwartz DA, Wurfel MM, Cambier JC. 2011. Identification and characterization of a loss-of-function human MPYS variant. *Genes and immunity* **12**(4): 263-269.

Kagan JC, Su T, Horng T, Chow A, Akira S, Medzhitov R. 2008. TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. *Nat Immunol* **9**(4): 361-368.

Kaiser WJ, Upton JW, Mocarski ES. 2008. Receptor-interacting protein homotypic interaction motif-dependent control of NF-kappa B activation via the DNA-dependent activator of IFN regulatory factors. *J Immunol* **181**(9): 6427-6434.

Karaghiosoff M, Neubauer H, Lassnig C, Kovarik P, Schindler H, Pircher H, McCoy B, Bogdan C, Decker T, Brem G et al. 2000. Partial impairment of cytokine responses in Tyk2-deficient mice. *Immunity* **13**(4): 549-560.

Kaser A, Lee AH, Franke A, Glickman JN, Zeissig S, Tilg H, Nieuwenhuis EE, Higgins DE, Schreiber S, Glimcher LH et al. 2008. XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell* **134**(5): 743-756.

Kato H, Sato S, Yoneyama M, Yamamoto M, Uematsu S, Matsui K, Tsujimura T, Takeda K, Fujita T, Takeuchi O et al. 2005. Cell type-specific involvement of RIG-I in antiviral response. *Immunity* **23**(1): 19-28.

Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, Uematsu S, Jung A, Kawai T, Ishii KJ et al. 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**(7089): 101-105.

Kawai T, Akira S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol* **11**(5): 373-384.

Kawai T, Takahashi K, Sato S, Coban C, Kumar H, Kato H, Ishii KJ, Takeuchi O, Akira S. 2005. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nature immunology* **6**(10): 981-988.

Kawane K, Fukuyama H, Kondoh G, Takeda J, Ohsawa Y, Uchiyama Y, Nagata S. 2001. Requirement of DNase II for definitive erythropoiesis in the mouse fetal liver. *Science* **292**(5521): 1546-1549.

Kawane K, Ohtani M, Miwa K, Kizawa T, Kanbara Y, Yoshioka Y, Yoshikawa H, Nagata S. 2006. Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages. *Nature* **443**(7114): 998-1002.

Kerur N, Veettil MV, Sharma-Walia N, Bottero V, Sadagopan S, Otageri P, Chandran B. 2011. IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection. *Cell Host Microbe* **9**(5): 363-375.

Kilic SS, Hacimustafaoglu M, Boisson-Dupuis S, Kreins AY, Grant AV, Abel L, Casanova JL. 2012. A patient with tyrosine kinase 2 deficiency without hyper-IgE syndrome. *The Journal of pediatrics* **160**(6): 1055-1057.

Kim YM, Brinkmann MM, Paquet ME, Ploegh HL. 2008. UNC93B1 delivers nucleotide-sensing toll-like receptors to endolysosomes. *Nature* **452**(7184): 234-238.

Klaman LD, Boss O, Peroni OD, Kim JK, Martino JL, Zabolotny JM, Moghal N, Lubkin M, Kim YB, Sharpe AH et al. 2000. Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. *Molecular and cellular biology* **20**(15): 5479-5489.

Krieg AM. 2002. CpG motifs in bacterial DNA and their immune effects. *Annual review of immunology* **20**: 709-760.

Krieg AM, Yi AK, Matson S, Waldschmidt TJ, Bishop GA, Teasdale R, Koretzky GA, Klinman DM. 1995. CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374**(6522): 546-549.

Kuhn R, Lohler J, Rennick D, Rajewsky K, Muller W. 1993. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* **75**(2): 263-274.

Kumar H, Kawai T, Akira S. 2011. Pathogen recognition by the innate immune system. *International reviews of immunology* **30**(1): 16-34.

Kumar H, Kawai T, Kato H, Sato S, Takahashi K, Coban C, Yamamoto M, Uematsu S, Ishii KJ, Takeuchi O et al. 2006. Essential role of IPS-1 in innate immune responses against RNA viruses. *The Journal of experimental medicine* **203**(7): 1795-1803.

Laguette N, Benkirane M. 2012. How SAMHD1 changes our view of viral restriction. *Trends Immunol* **33**(1): 26-33.

Lande R, Gregorio J, Facchinetti V, Chatterjee B, Wang YH, Homey B, Cao W, Wang YH, Su B, Nestle FO et al. 2007. Plasmacytoid dendritic cells sense self-DNA coupled with antimicrobial peptide. *Nature* **449**(7162): 564-569.

Latz E, Schoenemeyer A, Visintin A, Fitzgerald KA, Monks BG, Knetter CF, Lien E, Nilsen NJ, Espevik T, Golenbock DT. 2004. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nat Immunol* **5**(2): 190-198.

Le Goffic R, Balloy V, Lagranderie M, Alexopoulou L, Escriou N, Flavell R, Chignard M, Si-Tahar M. 2006. Detrimental contribution of the Toll-like receptor (TLR)3 to influenza A virus-induced acute pneumonia. *PLoS pathogens* **2**(6): e53.

Leadbetter EA, Rifkin IR, Hohlbaum AM, Beaudette BC, Shlomchik MJ, Marshak-Rothstein A. 2002. Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* **416**(6881): 603-607.

Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A. 2000. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. *Science* **289**(5488): 2350-2354.

Lee-Kirsch MA, Chowdhury D, Harvey S, Gong M, Senenko L, Engel K, Pfeiffer C, Hollis T, Gahr M, Perrino FW et al. 2007a. A mutation in TREX1 that impairs susceptibility to granzyme A-mediated cell death underlies familial chilblain lupus. *J Mol Med (Berl)* **85**(5): 531-537.

Lee-Kirsch MA, Gong M, Chowdhury D, Senenko L, Engel K, Lee YA, de Silva U, Bailey SL, Witte T, Vyse TJ et al. 2007b. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. *Nat Genet* **39**(9): 1065-1067.

Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. 1996. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell* **86**(6): 973-983.

Li S, Wang L, Berman M, Kong YY, Dorf ME. 2011. Mapping a dynamic innate immunity protein interaction network regulating type I interferon production. *Immunity* **35**(3): 426-440.

Li T, Diner BA, Chen J, Cristea IM. 2012. Acetylation modulates cellular distribution and DNA sensing ability of interferon-inducible protein IFI16. *Proceedings of the National Academy of Sciences of the United States of America* **109**(26): 10558-10563.

Loo YM, Fornek J, Crochet N, Bajwa G, Perwitasari O, Martinez-Sobrido L, Akira S, Gill MA, Garcia-Sastre A, Katze MG et al. 2008. Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *Journal of virology* **82**(1): 335-345.

Manry J, Laval G, Patin E, Fornarino S, Itan Y, Fumagalli M, Sironi M, Tichit M, Bouchier C, Casanova JL et al. 2011. Evolutionary genetic dissection of human interferons. *The Journal of experimental medicine* **208**(13): 2747-2759.

Marshak-Rothstein A, Rifkin IR. 2007. Immunologically active autoantigens: the role of toll-like receptors in the development of chronic inflammatory disease. *Annual review of immunology* **25**: 419-441.

Martinon F, Petrilli V, Mayor A, Tardivel A, Tschopp J. 2006. Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* **440**(7081): 237-241.

Masters SL, Simon A, Aksentijevich I, Kastner DL. 2009. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease (*). *Annual review of immunology* **27**: 621-668.

McWhirter SM, Fitzgerald KA, Rosains J, Rowe DC, Golenbock DT, Maniatis T. 2004. IFN-regulatory factor 3-dependent gene expression is defective in Tbk1-deficient mouse embryonic fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* **101**(1): 233-238.

Medzhitov R, Preston-Hurlburt P, Janeway CA, Jr. 1997. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* **388**(6640): 394-397.

Mellman I, Steinman RM. 2001. Dendritic cells: specialized and regulated antigen processing machines. *Cell* **106**(3): 255-258.

Melo MB, Kasperkovitz P, Cerny A, Konen-Waisman S, Kurt-Jones EA, Lien E, Beutler B, Howard JC, Golenbock DT, Gazzinelli RT. 2010. UNC93B1 mediates host resistance to infection with Toxoplasma gondii. *PLoS pathogens* **6**(8): e1001071.

Meraz MA, White JM, Sheehan KC, Bach EA, Rodig SJ, Dighe AS, Kaplan DH, Riley JK, Greenlund AC, Campbell D et al. 1996. Targeted disruption of the Stat1 gene in mice reveals unexpected physiologic specificity in the JAK-STAT signaling pathway. *Cell* **84**(3): 431-442.

Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, Tschopp J. 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* **437**(7062): 1167-1172.

Mibayashi M, Martinez-Sobrido L, Loo YM, Cardenas WB, Gale M, Jr., Garcia-Sastre A. 2007. Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. *Journal of virology* **81**(2): 514-524.

Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK et al. 2006. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**(6): 1283-1298.

Montgomery SB, Dermitzakis ET. 2011. From expression QTLs to personalized transcriptomics. *Nature reviews Genetics* **12**(4): 277-282.

Muller M, Briscoe J, Laxton C, Guschin D, Ziemiecki A, Silvennoinen O, Harpur AG, Barbieri G, Witthuhn BA, Schindler C et al. 1993. The protein tyrosine kinase JAK1 complements defects in interferon-alpha/beta and -gamma signal transduction. *Nature* **366**(6451): 129-135.

Muller U, Steinhoff U, Reis LF, Hemmi S, Pavlovic J, Zinkernagel RM, Aguet M. 1994. Functional role of type I and type II interferons in antiviral defense. *Science* **264**(5167): 1918-1921.

Muruve DA, Petrilli V, Zaiss AK, White LR, Clark SA, Ross PJ, Parks RJ, Tschopp J. 2008. The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response. *Nature* **452**(7183): 103-107.

Myers MP, Andersen JN, Cheng A, Tremblay ML, Horvath CM, Parisien JP, Salmeen A, Barford D, Tonks NK. 2001. TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B. *J Biol Chem* **276**(51): 47771-47774.

Napirei M, Karsunky H, Zevnik B, Stephan H, Mannherz HG, Moroy T. 2000. Features of systemic lupus erythematosus in Dnase1-deficient mice. *Nature genetics* **25**(2): 177-181.

Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. 2009. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* **324**(5925): 387-389.

Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, Bernards R. 2005. A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**(5): 773-786.

Novick D, Cohen B, Rubinstein M. 1994. The human interferon alpha/beta receptor: characterization and molecular cloning. *Cell* **77**(3): 391-400.

Okabe Y, Sano T, Nagata S. 2009. Regulation of the innate immune response by threonine-phosphatase of Eyes absent. *Nature* **460**(7254): 520-524.

Okada Y, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Kawaguchi T, Stahl EA, Kurreeman FA, Nishida N et al. 2012. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat Genet* **44**(5): 511-516.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**(5): 376-386.

Orzalli MH, Deluca NA, Knipe DM. 2012. Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein. *Proceedings of the National Academy of Sciences of the United States of America*.

Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H, Kobzik L, Higgins DE, Daly MJ, Bloom BR, Kramnik I. 2005. Ipr1 gene mediates innate immunity to tuberculosis. *Nature* **434**(7034): 767-772.

Paytubi S, Wang X, Lam YW, Izquierdo L, Hunter MJ, Jan E, Hundal HS, Proud CG. 2009. ABC50 promotes translation initiation in mammalian cells. *J Biol Chem* **284**(36): 24061-24073.

Perez de Diego R, Sancho-Shimizu V, Lorenzo L, Puel A, Plancoulaine S, Picard C, Herman M, Cardon A, Durandy A, Bustamante J et al. 2010. Human TRAF3 adaptor molecule deficiency leads to impaired Toll-like receptor 3 response and susceptibility to herpes simplex encephalitis. *Immunity* **33**(3): 400-411.

Picard C, Puel A, Bonnet M, Ku CL, Bustamante J, Yang K, Soudais C, Dupuis S, Feinberg J, Fieschi C et al. 2003. Pyogenic bacterial infections in humans with IRAK-4 deficiency. *Science* **299**(5615): 2076-2079.

Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A et al. 2012. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature*.

Pisitkun P, Deane JA, Difilippantonio MJ, Tarasenko T, Satterthwaite AB, Bolland S. 2006. Autoreactive B cell responses to RNA-related antigens due to TLR7 gene duplication. *Science* **312**(5780): 1669-1672.

Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M et al. 2007. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat Genet* **39**(12): 1477-1482.

Poltorak A, He X, Smirnova I, Liu MY, Van Huffel C, Du X, Birdwell D, Alejos E, Silva M, Galanos C et al. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**(5396): 2085-2088.

Quintana-Murci L, Alcais A, Abel L, Casanova JL. 2007. Immunology in natura: clinical, epidemiological and evolutionary genetics of infectious diseases. *Nat Immunol* **8**(11): 1165-1171.

Rakoff-Nahoum S, Hao L, Medzhitov R. 2006. Role of toll-like receptors in spontaneous commensal-dependent colitis. *Immunity* **25**(2): 319-329.

Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* **2**(8): 1896-1906.

Rebsamen M, Heinz LX, Meylan E, Michallet MC, Schroder K, Hofmann K, Vazquez J, Benedict CA, Tschopp J. 2009. DAI/ZBP1 recruits RIP1 and RIP3 through RIP homotypic interaction motifs to activate NF-kappaB. *EMBO Rep* **10**(8): 916-922.

Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA et al. 2009. Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. *Nature genetics* **41**(7): 829-832.

Rice GI, Kasher PR, Forte GM, Mannion NM, Greenwood SM, Szynkiewicz M, Dickerson JE, Bhaskar SS, Zampini M, Briggs TA et al. 2012. Mutations in ADAR1 cause Aicardi-Goutieres syndrome associated with a type I interferon signature. *Nat Genet*.

Roberts TL, Idris A, Dunn JA, Kelly GM, Burnton CM, Hodgson S, Hardy LL, Garceau V, Sweet MJ, Ross IL et al. 2009. HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA. *Science* **323**(5917): 1057-1060.

Rodig SJ, Meraz MA, White JM, Lampe PA, Riley JK, Arthur CD, King KL, Sheehan KC, Yin L, Pennica D et al. 1998. Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell* **93**(3): 373-383.

Romanoski CE, Lee S, Kim MJ, Ingram-Drake L, Plaisier CL, Yordanova R, Tilford C, Guan B, He A, Gargalovic PS et al. 2010. Systems genetics analysis of gene-by-environment interactions in human cells. *Am J Hum Genet* **86**(3): 399-410.

Rotem Z, Cox RA, Isaacs A. 1963. Inhibition of virus multiplication by foreign nucleic acid. *Nature* **197**: 564-566.

Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, Grace M, Dricot A, Askenazi M, Tavares M et al. 2012. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature*.

Rudolph D, Yeh WC, Wakeham A, Rudolph B, Nallainathan D, Potter J, Elia AJ, Mak TW. 2000. Severe liver degeneration and lack of NF-kappaB activation in NEMO/IKKgamma-deficient mice. *Genes & development* **14**(7): 854-862.

Samaniego LA, Neiderhiser L, DeLuca NA. 1998. Persistence and expression of the herpes simplex virus genome in the absence of immediate-early proteins. *Journal of virology* **72**(4): 3307-3320.

Sancho-Shimizu V, Perez de Diego R, Lorenzo L, Halwani R, Alangari A, Israelsson E, Fabrega S, Cardon A, Maluenda J, Tatematsu M et al. 2011. Herpes simplex encephalitis in children with autosomal recessive and dominant TRIF deficiency. *The Journal of clinical investigation* **121**(12): 4889-4902.

Satoh T, Kato H, Kumagai Y, Yoneyama M, Sato S, Matsushita K, Tsujimura T, Fujita T, Akira S, Takeuchi O. 2010. LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proceedings of the National Academy of Sciences of the United States of America* **107**(4): 1512-1517.

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G et al. 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* **422**(6929): 297-302.

Schindler C, Fu XY, Improta T, Aebersold R, Darnell JE, Jr. 1992. Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISGF-3 proteins that are activated by interferon alpha. *Proceedings of the National Academy of Sciences of the United States of America* **89**(16): 7836-7839.

Schroder K, Tschopp J. 2010. The inflammasomes. *Cell* **140**(6): 821-832.

Seth RB, Sun L, Ea CK, Chen ZJ. 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* **122**(5): 669-682.

Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, Gupta PB, Hao T, Silver SJ, Root DE et al. 2009. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139**(7): 1255-1267.

Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **1**(6): 2856-2860.

Shigemoto T, Kageyama M, Hirai R, Zheng J, Yoneyama M, Fujita T. 2009. Identification of loss of function mutations in human genes encoding RIG-I and MDA5: implications for resistance to type I diabetes. *J Biol Chem* **284**(20): 13348-13354.

Shimp SK, 3rd, Parson CD, Regna NL, Thomas AN, Chafin CB, Reilly CM, Nichole Rylander M. 2012. HSP90 inhibition by 17-DMAG reduces inflammation in J774 macrophages through suppression of Akt and nuclear factor-kappaB pathways. *Inflamm Res* **61**(5): 521-533.

Smirnov DA, Morley M, Shin E, Spielman RS, Cheung VG. 2009. Genetic analysis of radiation-induced changes in human gene expression. *Nature* **459**(7246): 587-591.

Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, Guja C, Ionescu-Tirgoviste C, Widmer B, Dunger DB et al. 2006. A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet* **38**(6): 617-619.

Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, Li Y, Kurreeman FA, Zhernakova A, Hinks A et al. 2010. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* **42**(6): 508-514.

Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD. 1998. How cells respond to interferons. *Annual review of biochemistry* **67**: 227-264.

Stetson DB, Ko JS, Heidmann T, Medzhitov R. 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* **134**(4): 587-598.

Stetson DB, Medzhitov R. 2006. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**(1): 93-103.

Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Akey JM. 2007. Gene-expression variation within and among human populations. *Am J Hum Genet* **80**(3): 502-509.

Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE, Sekowska M, Smith GD, Evans D, Gutierrez-Arcelus M et al. 2012. Patterns of cis regulatory variation in diverse human populations. *PLoS genetics* **8**(4): e1002639.

Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, Ingle CE, Dunning M, Flicek P, Koller D et al. 2007. Population genomics of human gene expression. *Nat Genet* **39**(10): 1217-1224.

Strobl B, Bubic I, Bruns U, Steinborn R, Lajko R, Kolbe T, Karaghiosoff M, Kalinke U, Jonjic S, Muller M. 2005. Novel functions of tyrosine kinase 2 in the antiviral defense against murine cytomegalovirus. *Journal of immunology* **175**(6): 4000-4008.

Sun Q, Sun L, Liu HH, Chen X, Seth RB, Forman J, Chen ZJ. 2006. The specific and essential role of MAVS in antiviral innate immune responses. *Immunity* **24**(5): 633-642.

Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E et al. 2009. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat Genet* **41**(10): 1100-1104.

Suzuki K, Mori A, Ishii KJ, Saito J, Singer DS, Klinman DM, Krause PR, Kohn LD. 1999. Activation of target-tissue immune-recognition molecules by double-stranded polynucleotides. *Proc Natl Acad Sci U S A* **96**(5): 2285-2290.

Suzuki N, Suzuki S, Duncan GS, Millar DG, Wada T, Mirtsos C, Takada H, Wakeham A, Itie A, Li S et al. 2002. Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature* **416**(6882): 750-756.

Tabeta K, Georgel P, Janssen E, Du X, Hoebe K, Crozat K, Mudd S, Shamel L, Sovath S, Goode J et al. 2004. Toll-like receptors 9 and 3 as essential components of innate immune defense against mouse cytomegalovirus infection. *Proceedings of the National Academy of Sciences of the United States of America* **101**(10): 3516-3521.

Tabeta K, Hoebe K, Janssen EM, Du X, Georgel P, Crozat K, Mudd S, Mann N, Sovath S, Goode J et al. 2006. The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. *Nat Immunol* **7**(2): 156-164.

Tager AM, Kradin RL, LaCamera P, Bercury SD, Campanella GS, Leary CP, Polosukhin V, Zhao LH, Sakamoto H, Blackwell TS et al. 2004. Inhibition of pulmonary fibrosis by the chemokine IP-10/CXCL10. *Am J Respir Cell Mol Biol* **31**(4): 395-404.

Takaoka A, Wang Z, Choi MK, Yanai H, Negishi H, Ban T, Lu Y, Miyagishi M, Kodama T, Honda K et al. 2007. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* **448**(7152): 501-505.

Takaoka A, Yanai H, Kondo S, Duncan G, Negishi H, Mizutani T, Kano S, Honda K, Ohba Y, Mak TW et al. 2005. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* **434**(7030): 243-249.

Takeuchi O, Akira S. 2010. Pattern recognition receptors and inflammation. *Cell* **140**(6): 805-820.

Tanaka Y, Chen ZJ. 2012. STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Sci Signal* **5**(214): ra20.

Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S et al. 2009. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* **41**(10): 1105-1109.

Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O'Huigin C, Kidd J, Kidd K, Khakoo SI, Alexander G et al. 2009. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* **461**(7265): 798-801.

Tsuchida T, Zou J, Saitoh T, Kumar H, Abe T, Matsuura Y, Kawai T, Akira S. 2010. The ubiquitin ligase TRIM56 regulates innate immune responses to intracellular double-stranded DNA. *Immunity* **33**(5): 765-776.

Turer EE, Tavares RM, Mortier E, Hitotsumatsu O, Advincula R, Lee B, Shifrin N, Malynn BA, Ma A. 2008. Homeostatic MyD88-dependent signals cause lethal inflamMation in the absence of A20. *The Journal of experimental medicine* **205**(2): 451-464.

Unterholzner L, Keating SE, Baran M, Horan KA, Jensen SB, Sharma S, Sirois CM, Jin T, Latz E, Xiao TS et al. 2010. IFI16 is an innate immune sensor for intracellular DNA. *Nature immunology* **11**(11): 997-1004.

Uze G, Lutfalla G, Gresser I. 1990. Genetic transfer of a functional human interferon alpha receptor into mouse cells: cloning and expression of its cDNA. *Cell* **60**(2): 225-234.

Vasseur E, Patin E, Laval G, Pajon S, Fornarino S, Crouau-Roy B, Quintana-Murci L. 2011. The selective footprints of viral pressures at the human RIG-I-like receptor family. *Human molecular genetics* **20**(22): 4462-4474.

Vazquez-Torres A, Vallance BA, Bergman MA, Finlay BB, Cookson BT, Jones-Carson J, Fang FC. 2004. Toll-like receptor 4 dependence of innate and adaptive immunity to Salmonella: importance of the Kupffer cell network. *Journal of immunology* **172**(10): 6202-6208.

Veals SA, Schindler C, Leonard D, Fu XY, Aebersold R, Darnell JE, Jr., Levy DE. 1992. Subunit of an alpha-interferon-responsive transcription factor is related to interferon regulatory factor and Myb families of DNA-binding proteins. *Molecular and cellular biology* **12**(8): 3315-3324.

Velazquez L, Fellous M, Stark GR, Pellegrini S. 1992. A protein tyrosine kinase in the interferon alpha/beta signaling pathway. *Cell* **70**(2): 313-322.

von Bernuth H, Picard C, Jin Z, Pankla R, Xiao H, Ku CL, Chrabieh M, Mustapha IB, Ghandil P, Camcioglu Y et al. 2008. Pyogenic bacterial infections in humans with MyD88 deficiency. *Science* **321**(5889): 691-696.

Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. 2004. Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nature medicine* **10**(12): 1366-1373.

Weiss DS, Raupach B, Takeda K, Akira S, Zychlinsky A. 2004. Toll-like receptors are temporally involved in host defense. *Journal of immunology* **172**(7): 4463-4469.

Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S, Wu P, Wiesmann C, Baker R, Boone DL et al. 2004. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature* **430**(7000): 694-699.

Wurfel MM, Park WY, Radella F, Ruzinski J, Sandstrom A, Strout J, Bumgarner RE, Martin TR. 2005. Identification of high and low responders to lipopolysaccharide in normal subjects: an unbiased approach to identify modulators of innate immunity. *Journal of immunology* **175**(4): 2570-2578.

Xu Y, Cheng G, Baltimore D. 1996. Targeted disruption of TRAF3 leads to postnatal lethality and defective T-dependent immune responses. *Immunity* **5**(5): 407-415.

Yan N, Cherepanov P, Daigle JE, Engelman A, Lieberman J. 2009. The SET complex acts as a barrier to autointegration of HIV-1. *PLoS pathogens* **5**(3): e1000327.

Yan N, O'Day E, Wheeler LA, Engelman A, Lieberman J. 2011. HIV DNA is heavily uracilated, which protects it from autointegration. *Proceedings of the National Academy of Sciences of the United States of America* **108**(22): 9244-9249.

Yan N, Regalado-Magdos AD, Stiggelbout B, Lee-Kirsch MA, Lieberman J. 2010. The cytosolic exonuclease TREX1 inhibits the innate immune response to human immunodeficiency virus type 1. *Nature immunology* **11**(11): 1005-1013.

Yanai H, Ban T, Wang Z, Choi MK, Kawamura T, Negishi H, Nakasato M, Lu Y, Hangai S, Koshiba R et al. 2009. HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. *Nature* **462**(7269): 99-103.

Yang P, An H, Liu X, Wen M, Zheng Y, Rui Y, Cao X. 2010. The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway. *Nature immunology* **11**(6): 487-494.

Yang YG, Lindahl T, Barnes DE. 2007. Trex1 exonuclease degrades ssDNA to prevent chronic checkpoint activation and autoimmune disease. *Cell* **131**(5): 873-886.

Yasutomo K, Horiuchi T, Kagami S, Tsukamoto H, Hashimura C, Urushihara M, Kuroda Y. 2001. Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat Genet* **28**(4): 313-314.

Yoneyama M, Kikuchi M, Matsumoto K, Imaizumi T, Miyagishi M, Taira K, Foy E, Loo YM, Gale M, Jr., Akira S et al. 2005. Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *Journal of immunology* **175**(5): 2851-2858.

Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T. 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5**(7): 730-737.

Yoshida H, Okabe Y, Kawane K, Fukuyama H, Nagata S. 2005. Lethal anemia caused by interferon-beta produced in mouse embryos carrying undigested DNA. *Nat Immunol* **6**(1): 49-56.

Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, Maouche S, Germain M, Lackner K, Rossmann H et al. 2010. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PloS one* **5**(5): e10693.

Zhang H, Cicchetti G, Onda H, Koon HB, Asrican K, Bajraszewski N, Vazquez F, Carpenter CL, Kwiatkowski DJ. 2003. Loss of Tsc1/Tsc2 activates mTOR and disrupts PI3K-Akt signaling through downregulation of PDGFR. *The Journal of clinical investigation* **112**(8): 1223-1233.

Zhang SY, Jouanguy E, Ugolini S, Smahi A, Elain G, Romero P, Segal D, Sancho-Shimizu V, Lorenzo L, Puel A et al. 2007. TLR3 deficiency in patients with herpes simplex encephalitis. *Science* **317**(5844): 1522-1527.

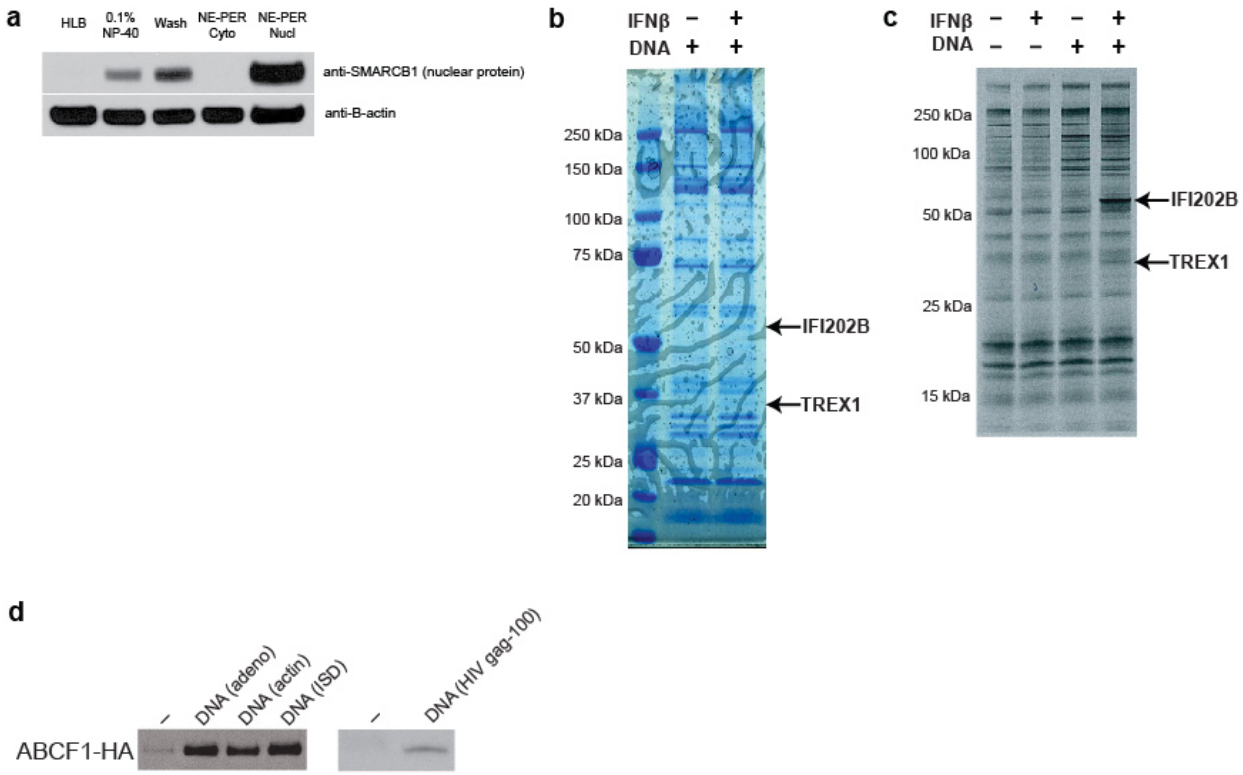
Zhang T, Hamza A, Cao X, Wang B, Yu S, Zhan CG, Sun D. 2008. A novel Hsp90 inhibitor to disrupt Hsp90/Cdc37 complex against pancreatic cancer cells. *Mol Cancer Ther* **7**(1): 162-170.

Zhang X, Brann TW, Zhou M, Yang J, Oguariri RM, Lidie KB, Imamichi H, Huang DW, Lempicki RA, Baseler MW et al. 2011a. Cutting edge: Ku70 is a novel cytosolic DNA sensor that induces type III rather than type I IFN. *Journal of immunology* **186**(8): 4541-4545.

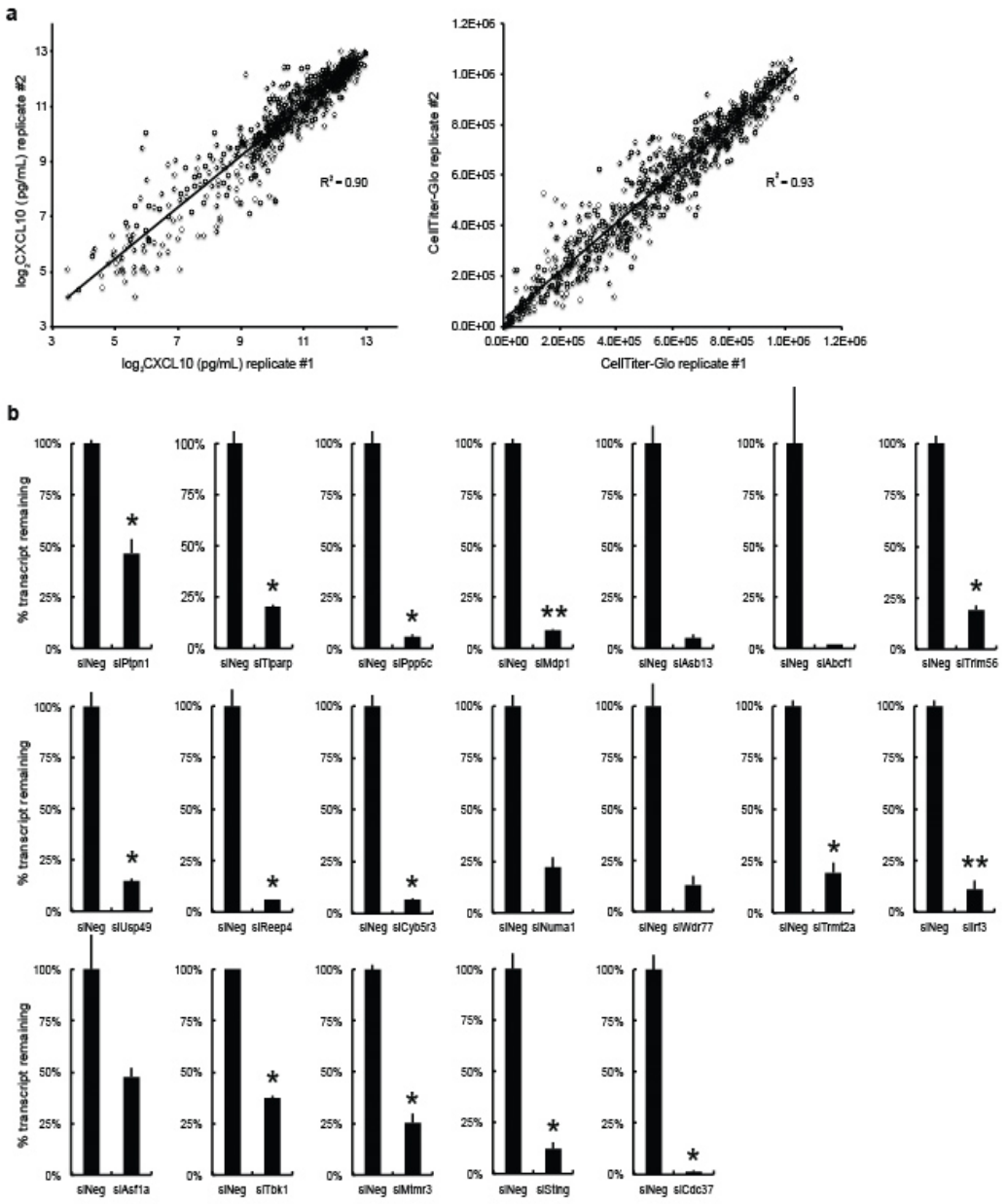
Zhang Z, Yuan B, Bao M, Lu N, Kim T, Liu YJ. 2011b. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nature immunology* **12**(10): 959-965.

Zheng L, Dai H, Zhou M, Li M, Singh P, Qiu J, Tsark W, Huang Q, Kernstine K, Zhang X et al. 2007. Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nature medicine* **13**(7): 812-819.

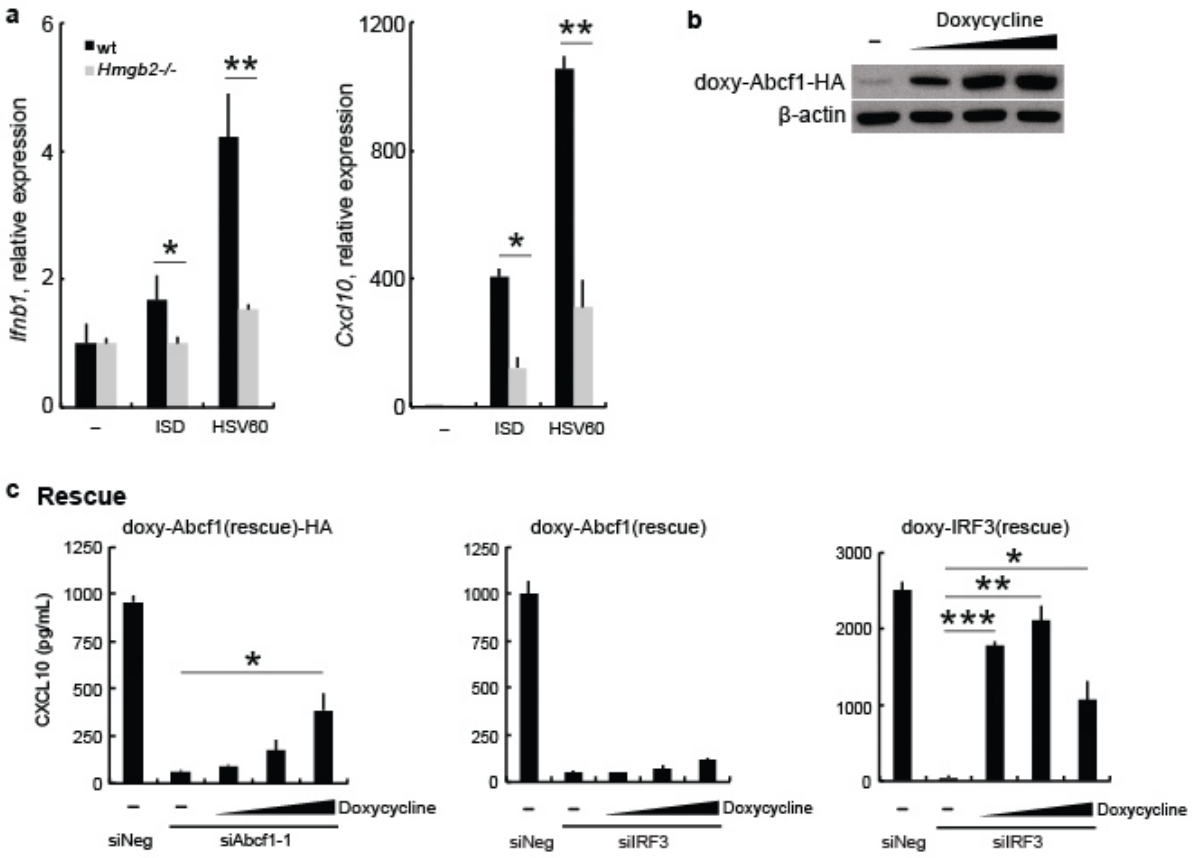
Appendix A: Supporting information for Chapter 1



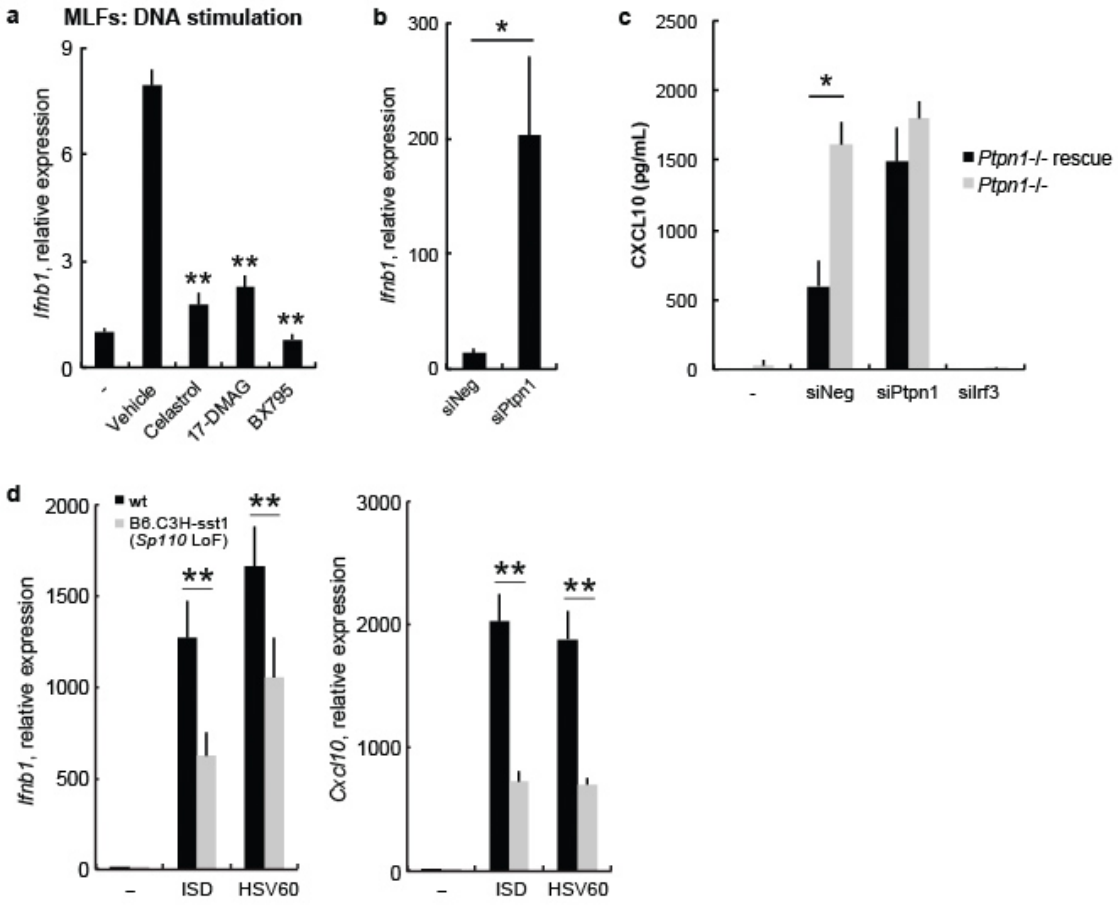
Supplementary Figure 1.1. Generation of a candidate gene set by curation and quantitative proteomics. (a) Cytoplasmic extract preparation. MEFs were lysed in panel of lysis buffers. Lysates were immunoblotted with anti-SMARCB1 antibody (nuclear protein) and with anti-β-actin antibody (loading) control. Lysis buffers are described in Supplementary Methods. (**b**) DNA pulldown. MEFs were pre-treated with IFNβ for 18 h or left unstimulated. Cytoplasmic extracts were prepared and incubated with biotinylated ISD. ISD-interacting proteins were precipitated with streptavidin beads, resolved by SDS-PAGE, and stained with Coomassie blue. Visualizable bands near 55 kDa and 35 kDa were identified by mass spectrometry as IFI202B and TREX1, respectively. (**c**) S35 DNA pulldown. MEFs were labeled with S35, and incubated with IFNβ for 6 h or left unstimulated. Cytoplasmic extracts were prepared and incubated with or without biotinylated ISD. ISD-interacting proteins were precipitated with streptavidin beads, resolved by SDS-PAGE, and visualized by autoradiography. (**d**) DNA pulldown assays were performed in MEFs stably expressing HA-tagged *Abcf1* with biotinylated 45-bp dsDNA of various sequences. Precipitates were resolved by SDS-PAGE and immunoblotted with anti-HA antibody.



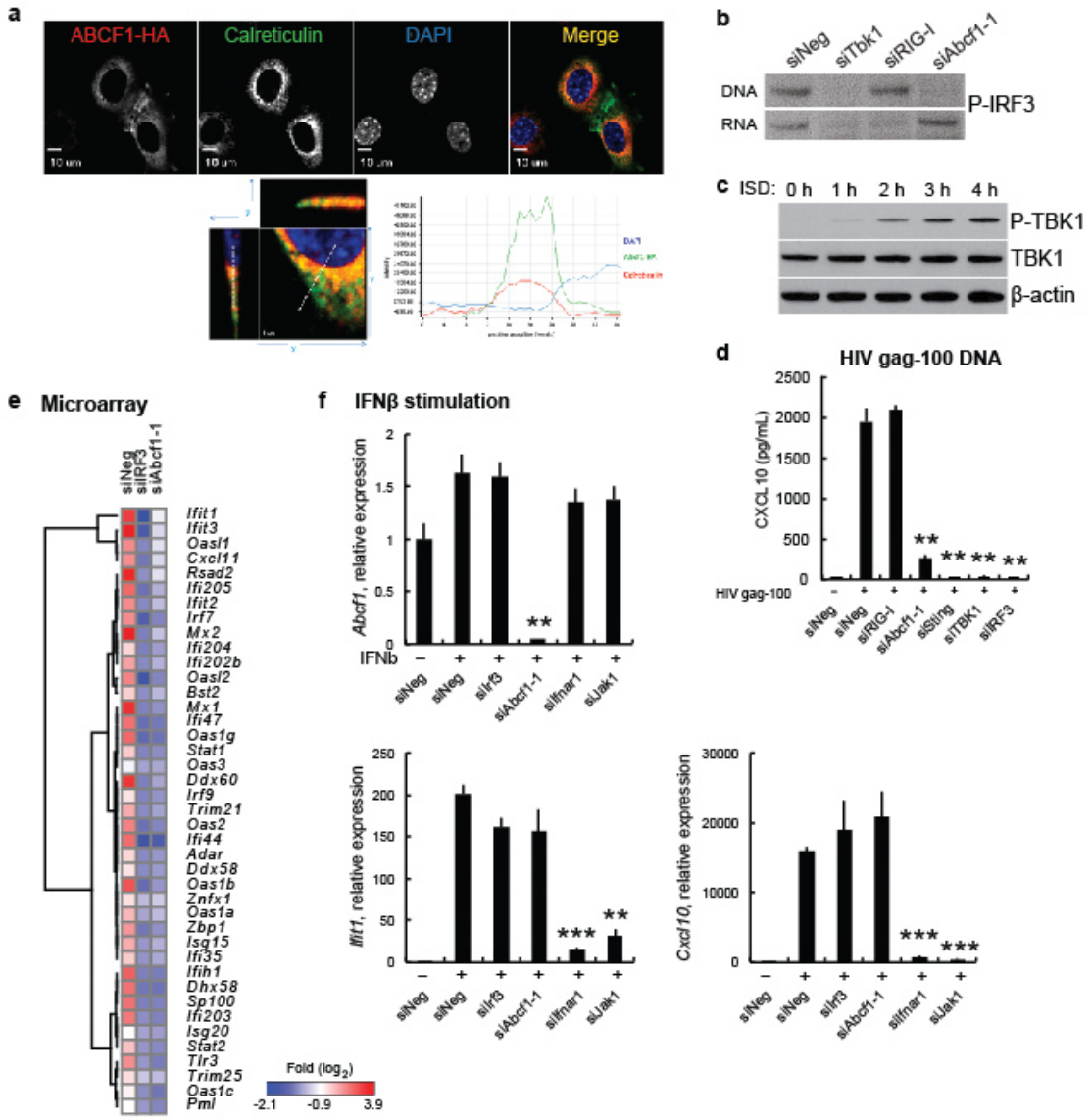
Supplementary Figure 1.2. High-throughput loss-of-function screening and network analysis of hits. (**a**) Robustness of siRNA screening assay. Log2 CXCL10 (pg/mL) and CellTiter-Glo luminescence levels from representative plates from the siRNA screen were graphed (x-axis) along with their respective values from replicate plates (y-axis). R^2 values are shown. (**b**) MEFs were treated with control siRNA (siNeg) or indicated siRNAs. Expression levels of respective genes was measured by qRT-PCR. Data is presented as mean and s.d. (n = 2). * $P < 0.05$, ** $P < 0.01$.



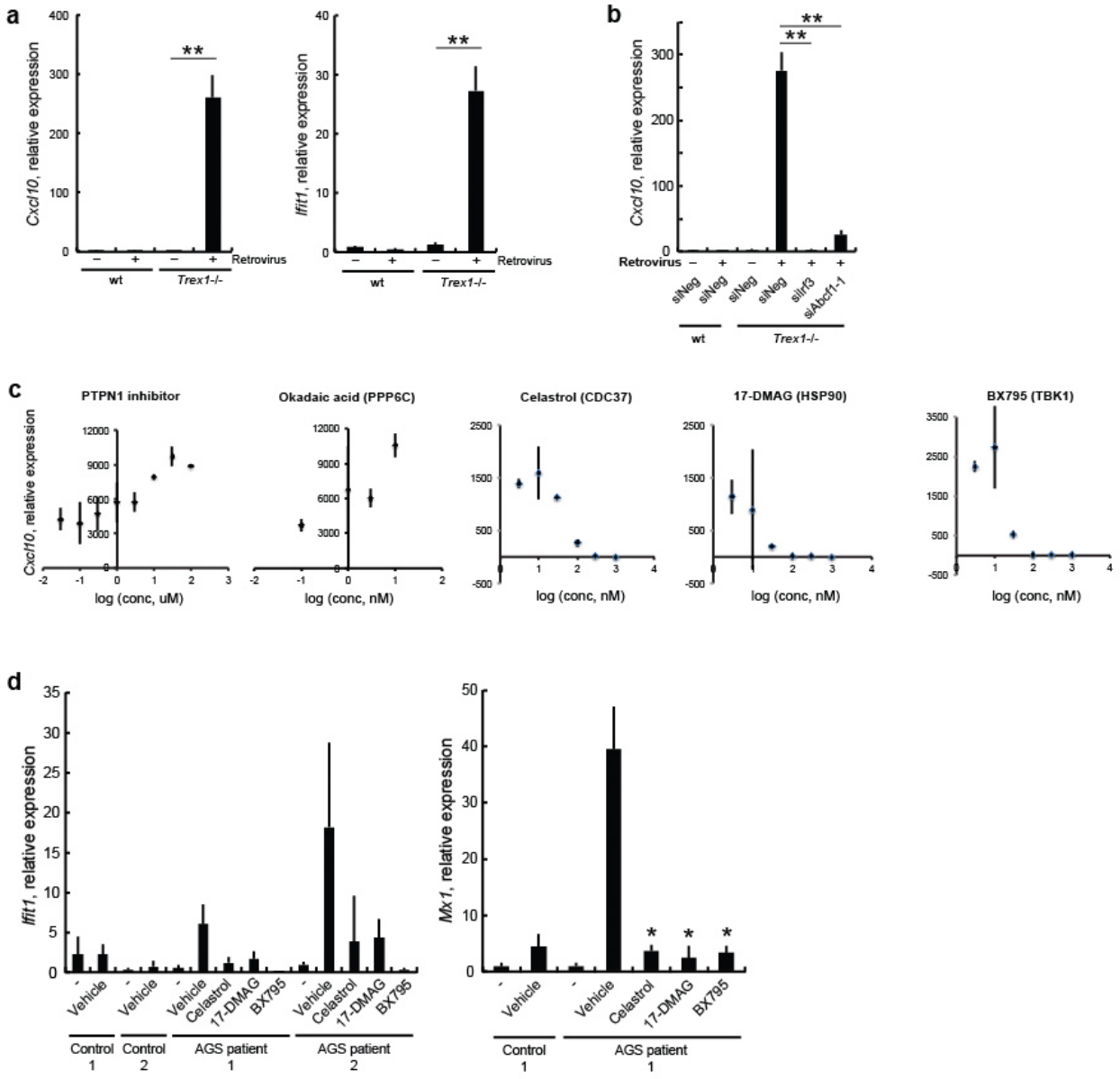
Supplementary Figure 1.3. Validation of screening hits. (**a**) *Hmgb2*-/- and wt MEFs were stimulated with DNA (ISD or HSV60 sequence) for 6 h, and *Cxcl10* induction was measured by qRT-PCR. * $P < 0.05$, ** $P < 0.01$. Data in all panels is presented as mean and s.d. (n = 3). (**b**) MEFs stably expressing doxycycline-inducible Abcf1-HA were treated with 0, 0.3, 3, or 30 ug/mL doxycycline. Lysates were immunoblotted with anti-HA or anti-β-actin antibody. (**c**) Validation by cDNA rescue. MEFs stably expressing doxycycline-inducible Abcf1(rescue)-HA, Abcf1(rescue), or Irf3(rescue) cDNA were treated with siNeg, siAbcf1, or siIrf3 along with doxycycline (0, 0.3, 3, 30 ug/mL). Cells were stimulated with ISD and CXCL10 production was measured by ELISA. * $P < 0.05$ compared with knockdown cells without doxycycline treatment.



Supplementary Figure 1.4. Targeting of screening hits by small molecule inhibitors. (**a**) Primary murine lung fibroblasts (MLFs) treated with small molecule inhibitors (500 nM celastrol, 750 nM 17-DMAG, 500 nM BX795) were stimulated with 4 ug/mL DNA (HIV gag-100) for 5.5 h, and induction levels of *Ifnb1* mRNA were determined by qRT-PCR. Data in all panels is presented as mean and s.d. (n = 3). ** $P < 0.01$ compared with DNA-stimulated vehicle-treated cells. (**b**) *Trex1*-/- MEFs treated with indicated siRNAs were stimulated with DNA (HIV gag-100) for 5.5 h, and *Ifnb1* induction was measured by qRT-PCR. * $P < 0.05$. (**c**) *Ptpn1*-/- MEFs and *Ptpn1*-/- MEFs rescued with *Ptpn1* cDNA were treated with indicated siRNAs and then stimulated with ISD for 26 h; CXCL10 levels were determined by ELISA. As expected, siPtpn1 increases CXCL10 levels in rescued MEFs, but does not have a significant phenotype in *Ptpn1*-/- MEFs. * $P < 0.05$. (**d**) B6.C3H-sst1 (Sp110 LoF) cDCs and wt control cells were stimulated with DNA for 6 h, and *Ifnb1* and *Cxcl10* induction was measured by qRT-PCR. ** $P < 0.01$.



Supplementary Figure 1.5. Identification of components of DNA sensing complex.
(**a**) Immunofluorescent microscropy of MEFs expressing HA-tagged *Abcf1*, stained for HA, calreticulin, and DAPI. Merge of images as well as line intensity graph is shown. (**b**) MEFs treated with indicated siRNAs were stimulated for 3 h with ISD or in vitro transcribed RNA, and lysates were immunoblotted with anti-phopho-IRF3 (Ser396) antibody. (**c**) MEFs were stimulated for with ISD for indicated times, and lysates were immunoblotted with anti-phopho-TBK1 (Ser172) antibody. (**d**) MEFs treated with indicated siRNAs were stimulated with DNA (HIV gag-100) for 26 h, and CXCL10 production was measured by ELISA. ** $P < 0.01$ compared with DNA-stimulated control siRNA (siNeg)-treated cells. Data is presented as mean and s.d. (n = 3). (**e**) Microarray analysis of MEFs treated with siRNAs and stimulated with ISD for 5.5 h. Log_2 fold changes of selected genes relative to unstimulated siNeg-treated cells are displayed as a heat map. Data are averages of two biological replicates. (**f**) Raw data from Fig. 5h. *Trex1*-/- MEFs treated with indicated siRNAs were stimulated with 300 U/mL IFNβ for 8 h, and expression of *Abcf1* mRNA as well as induction levels of ISGs were determined by qRT-PCR. ** $P < 0.01$, *** $P < 0.001$ compared with IFNβ-stimulated siNeg-treated cells. Data is presented as mean and s.d. (n = 3).



Supplementary Figure 1.6. Inhibition of identified regulators by RNAi or small molecules modulates the innate immune response to retroviral infection. (**a**) wt or *Trex1*-/- MEFs were infected with retrovirus for 21.5 h or left uninfected, and induction levels of *Cxcl10* and *Ifit1* were determined by qRT-PCR. ** $P < 0.01$. Data is presented as mean and s.d. (n = 3). (**b**) wt or *Trex1*-/- MEFs treated with indicated siRNAs were infected with retrovirus for 21.5 h, and induction levels of Cxcl10 mRNA were determined by qRT-PCR. ** $P < 0.01$. Data presented as mean and s.d. (n = 3). (**c**) *Trex1*-/- MEFs treated with small molecules at various doses were infected with retrovirus for 21.5 h, and induction levels of *Cxcl10* mRNA were determined by qRT-PCR. Data presented as mean and s.d. (n = 2). (**d**) *TREX1* mutant (AGS patient 1: R114H/D201ins; AGS patient 2: R114H/R114H) human fibroblasts and healthy control fibroblasts were treated with vehicle alone or small molecule inhibitors and infected with retrovirus. Induction levels of the ISGs, *MX1* and *IFIT1*, were determined by qRT-PCR. Data are averages of triplicate wells. Small molecules were used at: 500 nM celastrol, 100 nM 17-DMAG, and 500 nM BX795. * $P < 0.05$ compared with respective cells that were treated with vehicle control and infected with retrovirus.

Supplementary Table 1.1. STING-interacting SILAC hits

Gene Name	GeneID	log(H/L)	p-value
Sting	72512	2.793	0.00E+00
Atp2a2	11938	2.456	0.00E+00
Atp1a3	232975	2.798	0.00E+00
Thada	240174	2.374	1.78E-15
Slc39a7	14977	2.174	2.97E-13
Atp2b1	67972	2.079	3.01E-12
D19Bwg1357e	52874	-1.919	2.42E-11
Gemin5	216766	1.988	2.44E-11
Btaf1	107182	1.941	7.07E-11
Nup205	70699	1.904	1.57E-10
Vdac2	22334	1.807	1.22E-09
Parp1	11545	1.795	1.57E-09
Ano10	102566	1.786	1.87E-09
Parp14	547253	1.772	2.50E-09
Fer1l3	226101	1.724	6.48E-09
Surf4	20932	1.722	6.83E-09
Acaca	107476	1.664	2.09E-08
Top2a	21973	1.625	4.38E-08
Pnpt1	71701	1.537	2.16E-07
Eif4g1	208643	1.518	3.00E-07
Nolc1	70769	-1.446	3.11E-07
Eif3g	53356	1.506	3.67E-07
Eprs	107508	1.475	6.23E-07
Edc4	234699	1.449	9.64E-07
Mospd2	76763	1.443	1.07E-06
Cad	69719	1.412	1.77E-06
Cltc	67300	1.397	2.27E-06
Chd4	107932	1.395	2.35E-06
Mtap4	17758	1.377	3.12E-06
Abcb7	11306	1.322	7.37E-06
Dnajc13	235567	1.281	1.36E-05
Xpo5	72322	1.264	1.76E-05
Smchd1	74355	1.257	1.96E-05
Ascc3l1	320632	1.255	2.00E-05
Phb2	12034	1.226	3.05E-05
Dnmt1	13433	1.220	3.32E-05
Numa1	101706	1.210	3.82E-05
Sox2	20674	1.200	4.41E-05
Gvin1	74558	1.189	5.09E-05
Atp1a1	11928	1.177	6.04E-05
Ssr1	107513	1.148	9.01E-05
Nmi	64685	1.141	9.86E-05
Tfrc	22042	1.121	1.27E-04
Pigs	276846	1.107	1.53E-04
Lig3	16882	1.103	1.61E-04
Cnot1	234594	1.095	1.80E-04
Arbp	11837	1.091	1.89E-04
Far1	67420	1.088	1.97E-04
Man2a1	17158	1.064	2.65E-04
Ifi35	70110	1.061	2.77E-04
Erlin2	244373	1.039	3.65E-04
Sec61a1	53421	1.027	4.19E-04
Itprip	414801	1.015	4.84E-04
Sacm1l	83493	1.015	4.84E-04
Scrib	105782	1.009	5.23E-04
Upf2	326622	0.997	6.03E-04
Iqgap1	29875	0.982	7.15E-04
Sart3	53890	0.975	7.81E-04
Atxn2l	233871	0.972	8.04E-04
Cdc42bpa	226751	0.969	8.31E-04
Prpf8	192159	0.967	8.54E-04
Rpn1	103963	0.953	9.94E-04
C230096C10Rik	230866	0.934	1.24E-03
Glrx3	30926	0.930	1.29E-03
Abcd3	19299	0.927	1.34E-03
Ipo9	226432	0.913	1.56E-03
Mov10	17454	0.898	1.83E-03
Vars	22321	0.897	1.85E-03
Stt3a	16430	0.897	1.86E-03
Atp1b1	11931	0.884	2.13E-03
Ppp6r3	52036	0.883	2.16E-03
Ipo7	233726	0.876	2.32E-03
Gcn1l1	231659	0.875	2.34E-03
Rrbp1	81910	0.871	2.44E-03
Dhx37	208144	0.864	2.61E-03
Cope	59042	0.863	2.64E-03
Pfkl	18641	0.862	2.67E-03
Krt14	16664	-0.791	2.71E-03
Tmem33	67878	0.860	2.74E-03
Ddost	13200	0.858	2.79E-03
Parp9	80285	0.855	2.87E-03
Arhgef40	268739	0.846	3.17E-03
Mdn1	100019	0.834	3.56E-03
Sart1	20227	0.834	3.57E-03
Kif11	16551	0.832	3.65E-03
Ars2	83701	0.829	3.75E-03
Xpo7	65246	0.827	3.83E-03
Ipo5	70572	0.823	4.00E-03
Lmna	16905	-0.743	4.43E-03
Naa25	231713	0.801	4.93E-03
Mcm2	17216	0.801	4.94E-03
Nol5	55989	-0.728	5.11E-03
Hspa4	15525	0.796	5.22E-03
Mms19	72199	0.784	5.82E-03
Vdac3	22335	0.783	5.90E-03
Upf3b	68134	0.780	6.05E-03
Immt	76614	-0.705	6.39E-03
Upf1	19704	0.756	7.63E-03
Pnkp	59047	0.753	7.80E-03
Larp1	73158	0.748	8.22E-03
Dtx3l	209200	0.746	8.30E-03
S100a4	20198	0.737	9.06E-03
Ascc3	77987	0.735	9.21E-03
Rrp12	107094	0.728	9.80E-03
Thrap3	230753	0.727	9.92E-03
C330027C09Rik	224171	0.727	9.94E-03
Ubap2l	74383	0.724	1.01E-02
Rpn2	20014	0.722	1.04E-02
Smc2	14211	0.722	1.04E-02
Xpo1	103573	0.719	1.07E-02
Hyou1	12282	0.712	1.13E-02
Hsph1	15505	0.708	1.17E-02
Iars	105148	0.703	1.23E-02
Aimp2	231872	0.702	1.23E-02
Akr1b3	11677	0.701	1.25E-02
Slc25a3	18674	0.699	1.26E-02
Ddx56	52513	-0.623	1.34E-02
Clint1	216705	0.691	1.36E-02
Pelp1	75273	0.689	1.39E-02
Nisch	64652	0.687	1.41E-02
Vim	22352	-0.613	1.46E-02
Inf2	70435	0.682	1.47E-02
soat1	20652	0.654	1.86E-02
Hist1h1b	56702	0.654	1.86E-02
Bat2d	226562	0.654	1.86E-02
Cyb5r3	109754	0.647	1.97E-02
Casp3	12367	0.646	1.99E-02
Srprb	20818	0.641	2.07E-02
Nckap1l	105855	0.639	2.09E-02
Ldha	16828	0.634	2.19E-02
Tap1	21354	0.631	2.23E-02
Canx	12330	0.628	2.30E-02
Pfas	237823	0.626	2.33E-02
Ncapg	54392	0.616	2.52E-02
Smc3	13006	0.612	2.60E-02
Psmd1	70247	0.609	2.67E-02
Sec13	110379	0.602	2.81E-02
Fhod1	234686	0.599	2.89E-02
Lemd2	224640	-0.528	2.92E-02
Gart	14450	0.595	2.98E-02
Ncapd2	68298	0.594	3.00E-02
Samm50	68653	-0.523	3.02E-02
Mthfd1l	270685	0.588	3.13E-02
Dhx9	13211	0.582	3.27E-02
Slc3a2	17254	0.581	3.29E-02
Mat2a	232087	0.578	3.37E-02
Rab14	68365	0.578	3.38E-02
Mthfd2	17768	0.578	3.39E-02
Sec31a	69162	0.577	3.41E-02
Vdac1	22333	0.574	3.49E-02
Las1l	76130	0.569	3.60E-02
Acly	104112	0.565	3.71E-02
Sept9	53860	0.563	3.78E-02
Nomo1	211548	0.558	3.91E-02
Umps	22247	0.556	3.98E-02
Lars	107045	0.556	3.98E-02
Lig1	16881	0.333	4.19E-02
Aldoa	11674	0.547	4.24E-02
Wdr77	70465	0.544	4.32E-02
Usp15	14479	0.541	4.43E-02
Ddx58	230073	0.536	4.57E-02
Matr3	17184	0.536	4.59E-02
Ccar1	67500	0.531	4.75E-02
Rbm14	56275	0.528	4.85E-02
Slfn9	237886	0.526	4.93E-02

Supplementary Table 1.2. DNA-interacting SILAC hits.

Gene Name	GeneID	log(H/M)	log(H/L)
1110007L15Rik	67604	0.938	-0.092
1810029B16Rik	66282	0.595	-0.104
2810432D09Rik	69961	0.698	-0.154
8430406I07Rik	74528	1.040	-0.068
8430410A17Rik	232210	1.089	-0.152
A630055G03Rik	223970	1.309	0.054
Abcf1	224742	0.879	-0.076
Alkbh2	231642	0.826	-0.077
Anp32b	67628	1.405	-0.067
Anxa4	11746	1.134	-0.005
Anxa5	11747	1.317	-0.058
Apex1	11792	1.149	-0.058
Aplf	72103	0.894	-0.064
Aptx	66408	1.288	-0.100
Ascc1	69090	0.979	-0.095
Ascc2	75452	1.243	0.160
Ascc3	77987	1.180	0.177
Asf1a	66403	0.959	-0.091
Asf1b	66929	1.232	-0.149
Bst2	69550	0.322	1.186
Ccdc124	234388	0.616	-0.128
Cep170	545389	0.547	-0.302
Cggbp1	106143	0.553	-0.134
Cirbp	12696	1.016	-0.073
Cnn2	12798	1.383	-0.025
Crcp	12909	0.968	-0.001
Creb1	12912	1.078	0.023
Creb3l1	26427	1.234	-0.019
Creb3l2	208647	0.559	-0.193
Csda	56449	0.540	-0.152
Dazap1	70248	1.147	0.031
Ddb1	13194	0.864	-0.073
Ddb2	107986	1.085	-0.026
Ddx47	67755	0.327	-0.149
Ddx49	234374	0.450	-0.077
Dek	110052	1.090	-0.038
Dhx36	72162	0.476	-0.005
Dnajc9	108671	1.481	-0.066
Dr1	13486	1.461	-0.027
Drap1	66556	1.651	-0.028
Drg1	13494	0.426	-0.037
E430004N04Rik	210757	0.467	-1.653
Eef1d	66656	1.107	-0.039
Eif2s1	13665	0.511	-0.087
Eif2s2	67204	0.518	-0.077
Eif2s3x	26905	0.539	-0.082
Eif5b	226982	0.855	-0.066
Fam49b	223601	0.421	0.002
Fbxo18	50755	0.524	-0.076
Fen1	14156	1.241	-0.047
Frg1	14300	0.669	-0.117
Gata4	14463	0.661	0.032
Gnb2l1	14694	1.291	-0.030
H1fx	243529	0.606	-0.139
Hist1h1a	80838	1.561	-0.096
Hist1h1b	56702	1.580	-0.085
Hist1h1c	50708	1.460	-0.079
Hist1h1e	50709	1.008	-0.073
Hmga1	15361	1.444	-0.033
Hmga2	15364	1.603	-0.029
Hmgb1	15289	1.403	-0.081
Hmgb2	97165	1.550	-0.067
Hmgb3	15354	1.337	-0.132
Hmgn1	15312	1.044	-0.155
Hnrnpa0	77134	1.080	-0.008
Hnrnpa3	229279	1.346	-0.036
Hnrnpd	11991	1.316	-0.039
Hnrnpu	51810	1.083	-0.082
Hnrpab	15384	1.478	-0.035
Ifi202b	26388	1.267	1.171
Ifit1	15957	0.520	1.334
Kif22	110033	1.327	-0.059
Kif2a	16563	0.444	-0.075
Ldha	16828	1.559	-0.038
Ldhb	16832	1.408	-0.145
Lig3	16882	0.900	-0.100
LOC100044068	100044068	1.232	1.105
Mdh1	17449	1.641	-0.110
Mdh2	17448	1.435	-0.054
Morf4l2	56397	0.674	-0.069
Mpg	268395	1.082	-0.006
Msh2	17685	0.486	-0.094
Msh3	17686	0.848	-0.087
Msh6	17688	1.047	-0.233
Msi2	76626	0.821	0.026
Mtap1b	17755	1.118	-0.115
Mtap4	17758	1.025	-0.044
Mybbp1a	18432	1.100	-0.069
Nagk	56174	1.136	-0.069
Neil1	72774	1.131	-0.165
Nfib	18028	0.885	-0.228
Nthl1	18207	0.824	-0.078
Nuak2	74137	1.151	-0.105
Obfc1	108689	0.770	-0.101
Ogg1	18294	1.341	-0.087
Parp1	11545	1.184	-0.134
Parp2	11546	1.270	-0.082
Parp3	235587	1.244	0.203
Pcna	18538	0.759	-0.027
Pdia3	14827	0.471	-0.029
Phf6	70998	1.008	-0.088
Pkm2	18746	1.216	-0.050
Pnkp	59047	1.320	-0.072
Polb	18970	1.149	-0.104
Pold2	18972	0.738	0.039
Pole3	59001	1.155	-0.081
Polr1c	20016	1.748	-0.052
Polr1d	20018	0.809	-0.057
Polr2e	66420	1.652	-0.033
Polr2h	245841	0.934	-0.060
Polr3a	218832	1.043	-0.071
Polr3b	70428	0.779	-0.065
Polr3c	74414	1.582	-0.053
Polr3d	67065	0.805	-0.063
Polr3e	26939	0.836	-0.067
Polr3f	70408	1.558	-0.080
Polr3g	67486	0.964	-0.098
Polr3h	78929	0.776	-0.130
Prkar2b	19088	0.431	-0.057
Prrx1	18933	1.479	0.092
Psip1	101739	1.193	-0.073
Pycr2	69051	1.054	-0.132
Rabggtb	19352	1.359	-0.065
Rbm28	68272	0.816	-0.020
Rbm39	170791	0.691	-0.029
Rbms1	56878	1.100	-0.008
Rbms2	56516	1.483	0.051
Rbms3	207181	1.108	-0.157
Rbpj	19664	1.523	-0.076
Recql	19691	0.977	-0.081
Reep3	28193	1.211	-0.143
Reep4	72549	1.041	-0.098
Rfc1	19687	1.035	-0.097
Rfc2	19718	0.966	-0.082
Rfc3	69263	0.873	-0.100
Rfc4	106344	1.150	-0.088
Rfc5	72151	1.000	-0.077
Rpa1	68275	1.214	-0.070
Rpa2	19891	0.985	-0.066
Rpa3	68240	0.709	-0.077
Rpp38	227522	0.486	-0.135
Rrbp1	81910	1.133	-0.053
Rsl1d1	66409	0.556	-0.245
Samhd1	56045	0.641	0.092
Serbp1	66870	0.716	-0.049
Sfrs2	20382	0.534	-0.056
Sfrs3	20383	0.530	-0.066
Sfrs7	225027	0.494	0.021
Skp1a	21402	0.428	-0.082
Smarcal1	54380	0.880	-0.085
Snrpa	53607	0.805	-0.091
Snx9	66616	0.755	-0.067
Srm	20810	0.475	-0.114
Ssbp1	381760	0.909	-0.029
Sub1	20024	1.651	-0.052
Tbp	21374	1.209	-0.054
Tcfe3	209446	1.513	-0.057
Tcfeb	21425	1.900	-0.072
Tdg	21665	1.437	-0.178
Tdp1	104884	1.007	-0.095
Tead1	21676	0.493	-0.005
Tfam	21780	1.323	-0.031
Thex1	67276	0.616	-0.082
Thoc4	21681	0.546	-0.053
Thyn1	77862	1.408	-0.089
Top1	21969	1.210	-0.072
Top2a	21973	1.298	-0.004
Trex1	22040	1.932	0.932
Trip4	56404	1.375	0.124
Trmt6	66926	0.831	-0.004
Trmt61a	328162	1.334	-0.114
Tsnax	53424	1.528	-0.072
Ttf2	74044	1.627	0.026
Ubtf	21429	0.999	-0.084
Usp39	28035	0.676	-0.081
Vrk1	22367	1.042	-0.041
Wdr76	241627	0.618	0.003
Wrn	22427	1.471	0.012
Xbp1	22433	1.241	-0.156
Xpa	22590	1.439	-0.145
Xrcc1	22594	1.396	-0.148
Xrcc5	22596	1.113	-0.103
Xrcc6	14375	1.276	-0.100
Ybx1	22608	0.250	-0.083
Zc3h15	69082	0.610	-0.053
Znfx1	98999	1.835	0.404
Zranb3	226409	1.274	-0.126

Supplementary Table 1.3a. siRNA screen (DNA SILAC).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
1810029B16Rik	66282	898550	11.96
2810432D09Rik	69961	226980	9.54
8430406I07Rik	74528	324319	10.36
8430410A17Rik	232210	164207	10.24
Abcf1	224742	436318	6.17
Alkbh2	231642	781249	10.87
Anp32b	67628	891206	12.76
Anxa4	11746	594207	10.22
Anxa5	11747	642998	10.01
Apex1	11792	648863	12.28
Aplf	72103	612191	9.32
Aptx	66408	843595	12.79
Ascc1	69090	899815	12.15
Ascc2	75452	933005	12.40
Ascc3	77987	920689	11.85
Asf1a	66403	728637	9.13
Asf1b	66929	561790	9.59
Cggbp1	106143	292498	7.61
Cirbp	12696	411695	11.79
Cnn2	12798	161480	8.82
Crcp	12909	704860	10.48
Creb1	12912	561847	9.05
Creb3l1	26427	411838	10.54
Creb3l2	208647	995551	10.44
Dazap1	70248	479852	10.82
Ddb1	13194	188136	9.67
Ddb2	107986	764434	12.20
Dek	110052	820898	12.05
Dhx36	72162	255296	10.57
Dnajc9	108671	792280	11.19
Dr1	13486	807828	11.77
Drap1	66556	699328	10.55
Drg1	13494	96289	8.95
Eef1d	66656	601563	13.00
Eif2s1	13665	37353	5.01
Eif2s2	67204	22633	4.60
Eif2s3x	26905	63032	6.26
Eif5b	226982	89914	7.12
Fbxo18	50755	546232	11.97
Fen1	14156	557238	11.94
Frg1	14300	657963	10.48
Gata4	14463	850969	11.23
Get4	67604	752239	12.58
Gnb2l1	14694	122989	10.02
H1fx	243529	30898	5.10
Hist1h1a	80838	217293	11.03
Hist1h1b	56702	772476	12.97
Hist1h1c	50708	508069	12.71
Hist1h1e	50709	755248	12.30
Hmga1	15361	501887	9.85
Hmga2	15364	815652	11.81
Hmgb1	15289	767225	11.58
Hmgb2	97165	656372	7.31
Hmgb3	15354	830666	12.63
Hmgn1	15312	575799	11.69
Hnrnpu	51810	42304	5.51
Hnrpa3	229279	983795	12.25
Hnrpab	15384	532545	12.17
Hnrpd	11991	203355	6.31
Ifi202b	26388	101660	9.29
Ifit1	15957	813719	12.12
Kif22	110033	629769	11.52
Kif2a	16563	872593	10.38
Mdh1	17449	408462	11.83
Mdh2	17448	792701	12.74
Morf4l2	56397	868922	12.67
Mpg	268395	385051	9.86
Msh2	17685	830878	12.00
Msh3	17686	777511	12.31
Msh6	17688	710868	12.17
Msi2	76626	831157	12.83
Mtap1b	17755	342664	9.86
Mtap4	17758	567121	11.79
Mybbp1a	18432	107391	8.33
Nagk	56174	722064	12.23
Neil1	72774	698294	10.27
Nfib	18028	871568	11.86
Nthl1	18207	759335	11.85
Nuak2	74137	606236	12.74
Obfc1	108689	882905	11.64
Ogg1	18294	440781	10.75
Parp1	11545	696445	12.38
Parp2	11546	495252	11.67
Parp3	235587	420347	10.55
Pcna	18538	265036	10.68
Pdia3	14827	195208	10.13
Phf6	70998	693461	12.50
Pkm2	18746	351983	11.44
Pnkp	59047	829014	10.79
Polb	18970	905836	10.86
Pold2	18972	33411	9.00
Pole3	59001	766904	11.80
Polr3a	218832	63248	5.62
Polr3b	70428	266540	8.97
Polr3f	70408	53710	6.48
Prkar2b	19088	77542	6.56
Prrx1	18933	967833	11.89
Psip1	101739	431376	9.74
Pycr2	69051	747930	12.63
Rabggtb	19352	743104	11.56
Rbm28	68272	513202	9.70
Rbm39	170791	25132	5.29
Rbms1	56878	667531	11.40
Rbms2	56516	503962	11.74
Rbms3	207181	824174	11.80
Rbpj	19664	776597	11.99
Recql	19691	782063	11.67
Reep3	28193	752344	12.98
Reep4	72549	396200	6.52
Rfc1	19687	19021	5.84
Rfc4	106344	941082	11.99
Rpa1	68275	10061	5.05
Rpa2	19891	29230	6.22
Rpa3	68240	70375	6.92
Rrbp1	81910	661163	10.32
Rsl1d1	66409	283078	9.02
Samhd1	56045	253250	7.80
Serbp1	66870	971441	12.18
Sfrs2	20382	81501	6.19
Sfrs3	20383	71897	6.00
Sfrs7	225027	409102	10.61
Smarcal1	54380	285000	10.32
Snrpa	53607	489258	9.88
Snx9	66616	952324	12.92
Sub1	20024	927505	10.18
Tbp	21374	154417	9.07
Tcfe3	209446	826303	11.63
Tcfeb	21425	745699	12.20
Tdg	21665	455305	11.71
Tdp1	104884	429013	8.30
Tead1	21676	738063	12.58
Tfam	21780	761936	12.87
Themis	210757	672618	12.21
Thex1	67276	318535	9.88
Thoc4	21681	295545	11.14
Thyn1	77862	192873	9.56
Top1	21969	582200	10.24
Top2a	21973	556680	11.81
Trex1	22040	559668	12.20
Trip4	56404	609616	10.78
Trmt6	66926	767673	11.46
Trmt61a	328162	613226	12.79
Tsnax	53424	523860	12.76
Ttf2	74044	702655	12.83
Ubtf	21429	211731	10.16
Usp39	28035	321735	10.99
Vrk1	22367	746806	12.01
Wdr76	241627	451104	9.98
Wrn	22427	426305	10.52
Xbp1	22433	393908	10.32
Xpa	22590	720436	11.64
Xrcc1	22594	587468	10.24
Xrcc5	22596	718991	11.61
Xrcc6	14375	538894	10.46
Zc3h15	69082	783636	12.71
Znfx1	98999	876984	11.79

Supplementary Table 1.3b. siRNA screen (Microarray).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		945520	11.86
Irf3	54131	886771	8.66
Tbk1	56480	875356	8.25
Ddx41	72935	247041	10.46
Eya4	14051	802500.5	12.45
Mavs	228607	756700.1	10.54
Ikbke	56489	272150.9	9.77
Ikbkb	16150	744736.1	10.40
Tank	21353	553965.4	13.18
Ripk1	19766	462036.1	13.62
Jak1	16451	898386.1	10.02
Myd88	17874	545667.7	12.54
Trif	106759	253826.5	9.98
Nod2	257632	828969.3	11.80
1110003E01Rik	68552	622971	11.05
1110049F12Rik	66193	333550	11.56
1700010I14Rik	66931	254217	10.95
1700019G17Rik	75541	489101	10.39
4930547C10Rik	68274	629676	11.01
4933411K20Rik	66756	208430	9.84
9930111J21Rik	245240	695514	10.92
A230050P20Rik	319278	952951	11.94
Abhd11	68758	227566	6.27
Abhd3	106861	857853	11.60
Abtb2	99382	364495	12.01
Adar	56417	967911	12.43
Agrn	11603	762657	10.87
Agtrl1	23796	776544	12.02
Al451617	209387	860778	10.04
Aim1	11630	691219	10.63
Akt3	23797	651929	11.61
Alpk2	225638	734377	11.67
Apobec1	11810	666161	12.72
Apobec3	80287	575554	10.17
Arhgap9	216445	861944	11.44
Arid5a	214855	583754	12.48
Asb13	142688	712364	9.25
Asb14	142687	179894	8.41
Atp1b4	67821	655891	10.29
AY036118	170798	589310	11.08
B2m	12010	808280	11.03
Batf2	74481	358390	12.00
Baz2a	116848	238351	9.98
BC006779	229003	513478	12.93
Bptf	207165	779800	12.19
Brd4	57261	481951	10.32
Bst2	69550	618283	12.61
Btn1a1	12231	798103	12.31
Cabp1	29867	632222	11.51
Carhsp1	52502	340480	10.12
Casp7	12369	712238	12.42
Ccnb3	209091	812937	12.22
Cd274	60533	636840	11.80
Cd40	21939	596233	9.63
Cdadc1	71891	847505	11.66
Cenpj	219103	707149	10.59
Chek1	12649	589089	11.37
Chmp4c	66371	959464	10.68
Chmp5	76959	357976	12.01
Chn2	69993	909186	11.64
Clec2d	93694	860034	12.40
Cntn3	18488	480637	12.03
Copg2	54160	516824	11.79
Cops5	26754	246423	6.78
Cpeb3	208922	749199	12.56
Cphx	105594	687028	11.15
Csf1	12977	892342	11.71
Csprs	114564	175198	9.97
Cxcl10	15945	824505	7.84
D14Ertd668e	219132	372781	10.46
D1Pas1	110957	307301	10.08
Dab1	13131	800717	12.93
Daf2	13137	764737	11.42
Daxx	13163	711473	12.14
Dcpp1	13184	241912	7.30
Ddx3y	26900	249183	10.61
Ddx4	13206	513053	10.56
Ddx58	230073	166904	8.67
Dhx58	80861	378625	12.26
Ddx60	234311	862904	12.38
Dnase1l3	13421	330519	11.26
Ear5	54159	227534	9.96
Gm8995	668139	432088	10.10
Egfbp2	13647	861792	12.91
Eif2ak2	19106	188896	9.02
Elf1	13709	893660	11.70
Gm4836	22526	338428	11.65
Epm2aip1	77781	795583	12.52
Epsti1	108670	630676	11.02
Fam76b	72826	823445	12.92
Fas	14102	602820	12.72
Fbxl21	213311	538492	12.03
Foxa1	15375	770260	12.14
Galntl5	67909	209269	11.56
Gas7	14457	492498	9.64
Gbp1	14468	769162	11.45
Gbp2	14469	367108	8.57
Gbp3	55932	762919	10.84
Gbp4	17472	602249	10.97
Gbp5	229898	935426	12.88
Gbp6	229900	702127	11.99
Gbp9	236573	86616	6.62
Gca	227960	334541	10.15
Gch1	14528	511336	9.84
Gdap10	14546	681332	12.54
Gimap4	107526	255643	9.51
Glud1	14661	759506	12.81
Gngt1	14699	979035	12.09
Gpr137b	83924	713375	13.34
Gpr34	23890	671012	8.61
Gvin1	74558	774374	11.89
Hap1	15114	738386	12.45
Herc5	67138	509426	11.89
Higd1b	75689	737976	10.89
Hmg20a	66867	751720	11.20
Hmx3	15373	806384	13.13
Hopx	74318	272792	9.91
Hormad1	67981	194188	10.09
Hoxd13	15433	820280	11.57
Hsp90b1	22027	408000	11.72
Ifi203	15950	359386	8.55
Ifi204	15951	747272	12.51
Ifi205	226695	914170	11.91
Ifi27	76933	321026	10.55
Ifi35	70110	189528	9.93
Ifi44	99899	666608	11.47
Ifi47	15953	186496	7.67
Ifih1	71586	861028	12.22
Ifit2	15958	133398	10.10
Ifit3	15959	628079	12.64
Ifitm1	68713	582417	6.26
Ifitm2	80876	621922	12.22
Ifitm3	66141	200794	10.70
Igtp	16145	384972	9.53
Ihpk1	27399	916120	11.54
Iigp1	60440	563306	12.38
Iigp2	54396	846183	12.69
Il6	16193	756841	10.82
Insl6	27356	745690	10.24
Insm1	53626	323289	10.52
Irf2	16363	344555	8.83
Irf7	54123	677288	11.70
Irf9	16391	957040	10.60
Irgm	15944	645866	10.73
Isg20	57444	443497	11.96
Jam2	67374	786189	12.73
Jarid2	16468	804460	12.97
Kcne3	57442	753714	12.96
Klhl10	66720	907919	11.90
Klk1b16	16615	863107	11.63
Klk1b5	16622	519988	8.46
Klra16	27424	741652	10.89
Klra8	16639	401152	9.34
Lamp3	239739	991133	12.42
Lima1	65970	628132	12.55
Lin7b	22342	778454	11.73
Mcl1	17210	132281	8.35
Mef2c	17260	186496	9.23
Mela	17276	476356	10.41
Mov10	17454	506074	12.42
Mpa2l	100702	824552	13.33
Mpzl2	14012	855094	11.99
Mrps15	66407	555098	10.68
Ms4a4d	66607	984414	11.92
Mup4	17843	1039856	11.87
Mx1	17857	818464	11.40
Mx2	17858	299299	11.08
Mycl1	16918	456883	12.40
Myd88	17874	545668	12.54
Myef2	17876	1016931	12.10
Myt1	17932	766716	10.94
N4bp1	80750	451017	9.95
Nfxl1	100978	616138	9.82
Nlgn2	216856	514608	11.91
Nmi	64685	866355	11.64
Nphp3	74025	879664	12.41
Nr0b1	11614	908374	11.64
Nrtn	18188	553297	10.08
Ntng2	171171	778319	12.38
Nucks1	98415	824098	11.06
Nudt13	67725	883566	12.18
Numb	18222	826336	12.66
Oas1a	246730	841320	11.68
Oas1b	23961	610630	11.08
Oas1c	114643	643947	12.46
Oas1g	23960	875417	11.65
Oas2	246728	76748	6.91
Oas3	246727	199817	10.34
Oasl1	231655	629233	11.68
Oasl2	23962	287378	9.16
Ogfr	72075	667174	12.46
Olfm3	229759	489297	9.38
Oog1	193322	948235	12.34
Oplah	75475	967039	12.53
Pank2	74450	626055	11.84
Parp11	101187	415740	10.82
Parp12	243771	95965	8.17
Parp14	547253	898116	10.54
Parp9	80285	625718	12.79
Pax6	18508	667857	8.87
Pcbp2	18521	382768	9.62
Pbld2	67307	831332	12.64
Pdlim5	56376	610538	10.41
Pdx1	18609	406745	10.69
Peli1	67245	432051	9.91
Pex13	72129	707461	11.24
Phc3	241915	664454	10.04
Phex	18675	627895	12.90
Phf11	219131	421557	10.35
Phip	83946	764619	11.92
Phyhipl	70911	388043	11.90
Pigb	55981	1015601	12.57
Pik3c2g	18705	217498	11.39
Pla2g1b	18778	463113	13.08
Plac8	231507	879663	11.56
Plagl1	22634	300736	5.99
Plec1	18810	582278	11.50
Plekha4	69217	333320	11.16
Plk2	20620	11794	3.70
Plscr1	22038	536707	12.39
Plscr2	18828	598618	12.90
Pml	18854	442623	12.46
Pnp	18950	522037	11.75
Pnpla2	66853	587626	11.21
Pnpt1	71701	715127	12.36
Pnrc1	108767	751899	11.57
Podxl	27205	85720	8.72
Pou4f3	18998	280999	10.10
Ppm1a	19042	398060	11.67
Ppm1k	243382	704228	12.23
Prg2	19074	1000833	12.78
Psmb9	16912	783409	11.67
Psme1	19186	551584	12.97
Psrc1	56742	504843	11.39
Pvr	52118	762486	12.32
Pyhin1	236312	762328	12.71
Qk	19317	574878	12.34
Rab27a	11891	898937	12.07
Rab3c	67295	576793	8.86
Rai2	24004	93687	9.62
Rasgef1b	320292	848656	10.42
Rbck1	24105	410992	7.91
Rbm43	71684	898089	12.19
Rnf151	67504	486895	9.90
Rnf31	268749	524896	11.55
Rs1	20147	417205	9.61
Rsad2	58185	740467	12.42
Rshl1	83434	696178	12.62
Rsrc1	66880	568229	9.29
Rtp4	67775	637185	11.59
Samd9l	209086	839221	11.61
Sap30l	50724	225033	8.25
Scotin	66940	328723	8.37
Sec1	56546	804651	12.28
Serpina3m	20717	57360	4.99
Serpina6	12401	896298	12.08
Serpinb9	20723	707277	12.93
Serpini1	20713	848523	11.21
Shh	20423	503371	11.50
Slc1a3	20512	768230	12.85
Slc25a22	68267	271956	10.15
Slc25a28	246696	281271	10.14
Slc3a1	20532	635425	11.03
Slc6a14	56774	299739	9.82

Supplementary Table 1.3b. siRNA screen (Microarray, continued).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
Slc9a8	77031	970460	10.97
Slco1a6	28254	404338	11.05
Slfn10	237887	140896	8.56
Slfn2	20556	671520	11.99
Slfn3	20557	801240	11.25
Slfn4	20558	100099	6.92
Slfn5	327978	589349	12.32
Slfn8	276950	864810	11.47
Slfn9	237886	122982	6.11
Socs2	216233	812804	10.17
Sp100	20684	815062	11.87
Sp110	109032	676793	10.43
Sri	109552	988446	11.62
Stat1	20846	952376	10.51
Stat2	20847	267958	9.63
Stk31	77485	653389	9.49
Stk32c	57740	812730	12.38
Stk4	58231	430857	11.27
T2bp	211550	438109	8.74
Tap1	21354	597027	11.80
Tap2	21355	650915	10.72
Tapbp	21356	748280	11.22
Tbc1d10a	103724	641639	11.60
Tc2n	74413	773731	11.33
Tcirg1	27060	897895	11.78
Tcte3	21647	671322	11.00
Tdrd3	219249	308988	11.64
Tdrd7	100121	579683	11.34
Tgtp	21822	910768	12.61
Thrsp	21835	332374	10.47
Tiparp	99929	749564	13.93
Tlr3	142980	120136	7.85
Tmem106a	217203	231492	6.21
Tmem140	68487	928344	12.14
Tmem150c	231503	722485	11.82
Tmem229b	268567	651580	10.28
Tmem79	71913	838806	12.37
Tnc	21923	179956	11.75
Tnfaip3	21929	885004	11.87
Tnp1	21958	646111	11.30
Tor1aip1	208263	583996	11.49
Tor1aip2	240832	765239	10.43
Tor3a	30935	956014	12.38
Tra2a	101214	632183	11.24
Trafd1	231712	490427	10.49
Trim12c	319236	747450	11.08
Trim14	74735	991001	11.61
Trim21	20821	308798	11.26
Trim25	217069	458718	11.77
Trim26	22670	779900	11.15
Trim27	19720	804956	12.43
Trim30	20128	375426	10.00
Trim34	94094	834299	12.62
Trim56	384309	942177	9.80
Trim69	70928	617258	9.84
Twf1	19230	168922	5.57
Txndc10	67988	614666	10.31
Tyki	22169	973778	13.57
Ube1l	74153	687827	12.91
Ube2l6	56791	625045	11.91
Unc93b1	54445	969884	11.95
Usp18	24110	328952	7.40
Usp25	30940	348854	8.94
V1rd2	81016	944655	11.45
Vdac1	22333	651235	11.53
Vnn1	22361	187054	8.86
Vtcn1	242122	303170	11.14
Wdr78	242584	695988	12.39
Zbp1	58203	764439	11.56
Zc3hav1	78781	868569	11.81
Zfp313	81018	141227	11.17
Zkscan14	67235	899147	13.53

Supplementary Table 1.3c. siRNA screen (Protein-protein interactions).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)	Source
No siRNA		1047809	11.17	
Irf3	54131	943703	6.05	
Sting	72512	449277	4.87	
Abcb7	11306	285572	10.92	STING SILAC
Abcc1	17250	813911	9.97	STING IP (Li *et al*.)
Abcd3	19299	916831	11.13	STING SILAC
Acaca	107476	222596	11.19	STING SILAC
Acly	104112	751976	10.15	STING SILAC
Acsl3	74205	825343	11.04	STING IP (Li *et al*.)
Akr1b3	11677	489925	10.77	STING SILAC
Aldoa	11674	62551	6.95	STING SILAC
Arbp	11837	130169	7.68	STING SILAC
Ars2	83701	135128	8.52	STING SILAC
Atp1a1	11928	216473	9.63	STING SILAC
Atp1a3	232975	884686	11.46	STING SILAC
Atp1b1	11931	1007390	10.76	STING SILAC
Atp2a2	11938	181641	10.90	STING SILAC
Atxn2l	233871	200765	10.34	STING SILAC
Azi2	27215	211007	8.90	TBK1 IP (Li *et al*.)
C230096C10Rik	230866	547075	10.84	STING SILAC
C330027C09Rik	224171	665718	11.37	STING SILAC
Canx	12330	352520	11.15	STING SILAC
Casp3	12367	224600	9.84	STING SILAC
Ccar1	67500	155702	8.46	STING SILAC
Ccdc47	67163	161875	9.95	STING IP (Li *et al*.)
Cdc37	12539	591269	6.44	IKKi IP (Li *et al*.)
Cdc42bpg	240505	651000	11.16	TBK1 IP (Li *et al*.)
Clptm1l	218335	154572	7.28	STING IP (Li *et al*.)
Cnnm3	94218	604206	10.28	STING IP (Li *et al*.)
Cope	59042	146118	8.13	STING SILAC
Crebbp	12914	926537	11.16	IRF3 IP (Li *et al*.)
Cyb5r3	109754	808874	8.22	STING SILAC
D19Bwg1357e	52874	426976	10.60	STING SILAC
Ddost	13200	84854	6.72	STING IP (Li *et al*.)
Dnmt1	13433	775599	11.58	STING SILAC
Edc4	234699	74450	4.48	STING SILAC
Eif2c2	239528	936343	10.04	TBK1 IP (Li *et al*.)
Eif3g	53356	231503	9.62	STING SILAC
Ep300	328572	352174	9.38	IRF3 IP (Li *et al*.)
Eprs	107508	226688	9.76	STING SILAC
Erlin2	244373	483787	9.22	STING SILAC
Fhod1	234686	171043	5.49	STING SILAC
Fmr1	14265	978807	11.09	TBK1 IP (Li *et al*.)
Fxr2	23879	873097	9.64	TBK1 IP (Li *et al*.)
Gart	14450	275277	10.23	STING SILAC
Gemin5	216766	1071209	9.80	STING SILAC
Glmn	170823	728759	10.42	STING IP (Li *et al*.)
Glrx3	30926	401793	9.75	STING SILAC
H13	14950	506931	10.84	STING IP (Li *et al*.)
Hmmr	15366	822335	10.71	TBK1 IP (Li *et al*.)
Hsp110	15505	439823	11.16	STING SILAC
Hspa4	15525	276302	10.45	STING SILAC
Hyou1	12282	807701	10.87	STING SILAC
Iars	105148	91144	5.93	STING SILAC
Immt	76614	680691	10.96	STING SILAC
Inf2	70435	221239	11.04	STING SILAC
Ipo7	233726	670333	10.18	STING SILAC
Iqgap1	29875	467834	10.78	STING SILAC
Kif11	16551	112539	8.68	STING SILAC
Krt14	16664	889360	11.27	STING SILAC
Las1l	76130	139406	6.62	STING SILAC
Lemd2	224640	335175	9.62	STING SILAC
Leng4	77582	911706	11.21	STING IP (Li *et al*.)
Letm1	56384	267767	10.05	TBK1 IP (Li *et al*.)
Mat2a	232087	251881	9.69	STING SILAC
Mdn1	100019	228103	9.90	STING SILAC
Mib1	225164	897131	10.68	TBK1 IP (Li *et al*.)
Mlstd2	67420	846922	11.43	STING SILAC
Mms19	72199	958618	10.79	STING SILAC
Mthfd1l	270685	705985	11.04	STING SILAC
Mthfd2	17768	1079278	10.92	STING SILAC
Naa25	231713	127363	6.65	STING SILAC
Ncapd2	68298	205927	10.61	STING SILAC
Nckap1l	105855	858081	10.07	STING SILAC
Nisch	64652	643362	10.45	STING SILAC
Nol5	55989	458193	8.80	STING SILAC
Nolc1	70769	552997	9.91	STING SILAC
Nomo1	211548	1081120	10.87	STING SILAC
Npc1	18145	286666	10.93	STING IP (Li *et al*.)
Numa1	101706	704930	8.99	STING SILAC
Optn	71648	445097	11.47	TBK1 IP (Li *et al*.)
Papd1	67440	554579	10.91	TBK1 IP (Li *et al*.)
Pelp1	75273	968999	11.67	STING SILAC
Pfkl	18641	458490	9.97	STING SILAC
Phb2	12034	101658	6.89	STING SILAC
Pigs	276846	738487	10.28	STING SILAC
Rab14	68365	755449	11.09	STING SILAC
Ranbp5	70572	304861	11.31	STING SILAC

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)	Source
Rb1	19645	975497	10.26	IRF3 IP (Li *et al*.)
Rbbp5	213464	116282	8.75	TBK1 IP (Li *et al*.)
Rbl1	19650	861637	10.78	IRF3 IP (Li *et al*.)
Rbl2	19651	651901	10.89	IRF3 IP (Li *et al*.)
Rbm14	56275	174530	8.54	STING SILAC
Rpn1	103963	191481	8.81	STING SILAC
Rpn2	20014	529801	9.83	STING SILAC
Rrp12	107094	631721	10.96	STING SILAC
S100a4	20198	934813	11.60	STING SILAC
Sacm1l	83493	911915	10.87	STING SILAC
Samm50	68653	123391	6.83	STING SILAC
Saps3	52036	561837	11.21	STING SILAC
Sart1	20227	142917	7.79	STING SILAC
Scrib	105782	360909	8.76	STING SILAC
Sec13	110379	309654	9.77	STING SILAC
Sec61a1	53421	43241	4.71	STING IP (Li *et al*.)
Sept9	53860	445453	10.37	STING SILAC
Sgpl1	20397	530893	10.63	STING IP (Li *et al*.)
Slc25a3	18674	447242	9.06	STING SILAC
Slc39a14	213053	1145088	10.97	STING IP (Li *et al*.)
Slc39a7	14977	437607	10.12	STING SILAC
Slc3a2	17254	914803	10.94	STING SILAC
Smc2	14211	220883	10.11	STING SILAC
Smc3	13006	399156	9.49	STING SILAC
Smchd1	74355	159721	9.36	STING SILAC
Soat1	20652	972171	11.00	STING SILAC
Srprb	20818	720761	10.23	STING SILAC
Ssr1	107513	845359	10.33	STING SILAC
Stt3a	16430	384939	10.18	STING IP (Li *et al*.)
Surf4	20932	761951	10.33	STING IP (Li *et al*.)
Tbkbp1	73174	261117	10.22	TBK1 IP (Li *et al*.)
Tfrc	22042	248458	10.34	STING SILAC
Thada	240174	887220	11.19	STING SILAC
Thrap3	230753	740700	11.13	STING SILAC
Tmem16k	102566	662243	10.02	STING SILAC
Tmem33	67878	177005	7.75	STING SILAC
Txlna	109658	491961	11.11	TBK1 IP (Li *et al*.)
Txlng	353170	187343	9.55	TBK1 IP (Li *et al*.)
Ubap2l	74383	489682	10.00	STING SILAC
Umps	22247	357745	10.13	STING SILAC
Vars	22321	272203	9.88	STING SILAC
Vdac2	22334	290085	9.62	STING SILAC
Vdac3	22335	630456	10.39	STING SILAC
Vim	22352	711113	10.39	STING SILAC
Wdr77	70465	442113	7.42	STING SILAC
Xpo1	103573	331451	10.81	STING SILAC
Xpo5	72322	900258	11.33	STING SILAC
Xpo7	65246	769730	10.47	STING SILAC
Yipf5	67180	230973	10.20	STING IP (Li *et al*.)

Supplementary Table 1.3d. siRNA screen (Phosphatases).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		1252393	12.47
Irf3	54131	1152920	8.69
4921523A10Rik	110332	956200	10.38
Acp1	11431	511613	10.99
C79127	232941	707507	13.48
Cdc14a	229776	964373	12.22
Cdc14b	218294	1064040	13.85
Cdc25a	12530	995200	13.13
Cdc25b	12531	1144240	12.87
Cdc25c	12532	946507	13.60
Cdkn3	72391	1248293	12.33
Ctdp1	67655	827373	10.87
Ctdsp1	227292	1213600	12.22
Ctdsp2	52468	843507	14.12
Ctdspl	69274	1113800	9.43
Dnajc6	72685	1070400	12.04
Dullard	67181	753453	13.31
Dusp1	19252	735547	13.53
Dusp10	63953	1051373	12.64
Dusp11	72102	1228907	12.46
Dusp12	80915	938053	12.29
Dusp13	27389	1094867	10.76
Dusp14	56405	938120	12.66
Dusp15	252864	617880	10.72
Dusp18	75219	801053	12.21
Dusp19	68082	754120	11.12
Dusp2	13537	901267	12.26
Dusp22	105352	1243907	12.29
Dusp23	68440	601680	10.35
Dusp26	66959	912147	14.31
Dusp28	67446	318053	12.25
Dusp3	72349	693360	13.77
Dusp4	319520	776320	12.56
Dusp5	240672	1000960	13.64
Dusp6	67603	1007187	11.72
Dusp7	235584	558893	11.25
Dusp8	18218	905947	9.86
Dusp9	75590	1058827	12.06
Epm2a	13853	833600	10.26
Eya1	14048	1174227	12.18
Eya2	14049	752653	10.62
Eya3	14050	1020947	12.79
Ilkap	67444	652627	12.47
Map2k1	26395	1137360	13.11
Mdp1	67881	857213	14.70
Mtm1	17772	1183827	13.16
Mtmr1	53332	574280	11.28
Mtmr12	268783	968920	10.97
Mtmr2	77116	592600	14.27
Mtmr3	74302	986827	9.96
Mtmr4	170749	992373	14.04
Mtmr6	219135	767400	10.77
Mtmr7	54384	1359973	12.64
Phlpp	98432	896187	11.90
Phpt1	75454	303947	10.98
Ppef2	19023	457947	8.49
Ppm1b	19043	926893	12.59
Ppm1d	53892	810093	12.45
Ppm1f	68606	562760	11.53
Ppm1g	14208	784640	13.80
Ppm1h	319468	948480	12.91
Ppm1j	71887	715200	10.52
Ppm1l	242083	873053	11.96
Ppm1m	67905	795653	11.70
Ppp1ca	19045	527160	13.90
Ppp1cb	19046	510400	12.25
Ppp1cc	19047	935707	13.26
Ppp1r12a	17931	257280	11.13
Ppp1r15b	108954	938533	12.15
Ppp1r3a	140491	829880	12.70
Ppp2ca	19052	1102253	12.30
Ppp2cb	19053	893013	10.57
Ppp3ca	19055	1099787	12.84
Ppp3cb	19056	531053	12.68
Ppp3cc	19057	637547	10.72
Ppp3r2	19059	687747	12.54
Ppp4c	56420	952093	11.67
Ppp5c	19060	791413	11.09
Ppp6c	67857	1275907	14.76
Pptc7	320717	670973	11.47
Pten	19211	626493	13.55
Ptp4a1	19243	1054733	12.24
Ptp4a2	19244	1036307	10.91
Ptp4a3	19245	1060640	13.30
Ptpdc1	218232	945440	13.23
Ptpla	30963	963213	10.50
Ptplb	70757	1014240	12.70
Ptpmt1	66461	1015600	13.31

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
Ptpn1	19246	1171400	14.16
Ptpn11	19247	794520	11.80
Ptpn12	19248	593733	12.57
Ptpn13	19249	1017880	12.40
Ptpn14	19250	770053	10.57
Ptpn18	19253	791467	13.57
Ptpn2	19255	931493	13.07
Ptpn20	19256	269093	12.55
Ptpn21	24000	725027	12.65
Ptpn22	19260	806120	10.59
Ptpn23	104831	768240	13.00
Ptpn4	19258	734347	12.85
Ptpn5	19259	883427	12.17
Ptpn7	320139	924387	14.37
Ptpn9	56294	892093	11.82
Ptpra	19262	1048747	13.68
Ptprb	19263	836013	10.10
Ptprc	19264	1095280	12.50
Ptpre	19267	1201827	12.98
Ptprf	19268	918360	11.71
Ptprg	19270	584987	10.97
Ptprj	19271	809973	11.82
Ptprk	19272	1000120	13.19
Ptprm	19274	793320	12.39
Ptprn	19275	889227	12.82
Ptprn2	19276	826760	12.37
Ptprr	19279	710000	11.32
Ptprs	19280	725907	12.36
Ptprt	19281	1120840	11.90
Ptpru	19273	1004160	13.44
Ptprv	13924	744733	10.77
Rngtt	24018	615387	9.54
Ssh1	231637	867280	14.20
Ssh2	237860	798600	10.97
Ssh3	245857	744827	10.60
Ssu72	68991	793067	10.23
Styx	56291	776827	13.63
Tenc1	209039	340813	9.76
Timm50	66525	1096107	12.72
Ublcp1	79560	1219333	12.27

Supplementary Table 1.3e. siRNA screen (Deubiquitinases).

Gene Name	GeneID	CellTiter-Glo	log2(CXCL10, pg/mL)
No siRNA		1252393	12.47
Irf3	54131	1152920	8.69
Atxn3	110616	580653	12.34
Bap1	104416	822693	12.14
Brcc3	210766	1266227	12.73
Cops6	26893	825187	11.81
Cyld	74256	1109120	14.78
Dub1	13531	804600	13.31
Dub1a	381944	734987	12.16
Dub2	13532	823893	12.95
Eif3f	66085	622373	10.59
Eif3h	68135	1094093	13.41
Gm5136	368203	1241480	11.81
Josd1	74158	742213	12.97
Josd2	66124	1049867	11.46
Josd3	75316	721000	12.42
Mysm1	320713	886320	12.68
Otub1	107260	1079320	12.16
Otub2	68149	1092760	13.28
Otud5	54644	630227	10.94
Otud6b	72201	553160	13.46
Prpf8	192159	376427	8.85
Psmd14	59029	87347	6.71
Psmd7	17463	180093	7.78
Stambp	70527	1174547	13.99
Stambpl1	76630	1258400	13.72
Uchl1	22223	1234920	13.82
Uchl3	50933	1069520	11.68
Uchl4	93841	921360	12.41
Uchl5	56207	1271973	10.93
Usp1	230484	575520	11.12
Usp10	22224	798747	13.04
Usp11	236733	824427	11.23
Usp12	22217	845787	14.96
Usp13	72607	954680	12.26
Usp14	59025	896133	10.78
Usp15	14479	701933	11.68
Usp16	74112	914093	12.33
Usp2	53376	773147	11.25
Usp20	74270	765533	10.24
Usp21	30941	917680	11.75
Usp22	216825	782653	12.23
Usp24	329908	606640	9.49
Usp26	83563	993280	10.92
Usp27x	54651	577387	14.39
Usp28	235323	1052707	11.43
Usp29	57775	196280	10.29
Usp3	235441	654373	11.11
Usp30	100756	1084307	12.86
Usp31	76179	1241680	12.84
Usp33	170822	889720	12.55
Usp36	72344	1152133	12.80
Usp37	319651	1255640	13.62
Usp38	74841	1072520	9.46
Usp4	22258	568280	12.27
Usp42	76800	1053493	12.20
Usp43	216835	905573	9.28
Usp44	327799	1165787	13.01
Usp45	77593	1245093	11.95
Usp46	69727	921693	12.31
Usp47	74996	1053267	13.77
Usp48	170707	1030533	13.56
Usp49	224836	866147	9.80
Usp5	22225	485787	13.42
Usp50	75083	852840	10.87
Usp53	99526	713600	13.82
Usp54	78787	1012507	14.03
Usp7	252870	986733	10.74
Usp8	84092	697547	13.77
Usp9x	22284	1255720	12.83
Usp9y	107868	1033973	11.51
Vcpip1	70675	1002267	13.13
Yod1	226418	1103307	12.78

Supplementary Table 1.4. ABCF1-interacting SILAC hits.

Gene Name	GeneID	log(M/L)	log(H/L)	p-value M/L	p-value H/L
Abcf1	224742	2.962	3.587	0.00E+00	0.00E+00
Hmgb2	97165	1.046	0.962	8.71E-07	3.50E-03
Azi1	12009	0.972	0.061	4.44E-06	4.90E-01
Pdxdc1	94184	0.870	0.258	3.47E-05	2.84E-01
Tubal3	238463	0.864	1.236	3.86E-05	2.13E-04
Tbl2	27368	0.837	0.610	6.42E-05	5.13E-02
Slc1a5	20514	0.827	0.464	7.79E-05	1.16E-01
Iigp2	54396	0.825	0.460	7.95E-05	1.19E-01
Pfkl	18641	0.814	0.793	9.82E-05	1.44E-02
Vcam1	22329	0.785	0.298	1.63E-04	2.46E-01
Cd47	16423	0.775	0.901	1.97E-04	5.99E-03
Iigp1	60440	0.753	0.271	2.83E-04	2.71E-01
2310014H01Rik	76448	0.751	-0.019	2.92E-04	3.95E-01
Slfn9	237886	0.743	0.840	3.33E-04	9.97E-03
Itpr3	16440	0.726	0.301	4.47E-04	2.43E-01
Pla2g4a	18783	0.720	0.042	4.93E-04	4.67E-01
Tgtp	21822	0.691	0.416	7.79E-04	1.48E-01
Tap2	21355	0.672	1.055	1.04E-03	1.45E-03
Keap1	50868	0.652	-0.437	1.40E-03	6.29E-02
Ldha	16828	0.647	1.124	1.53E-03	7.22E-04
Ints3	229543	0.615	0.473	2.41E-03	1.11E-01
Gbp2	14469	0.606	0.893	2.74E-03	6.43E-03
Gls	14660	0.602	0.668	2.92E-03	3.54E-02
Prpf31	68988	0.600	0.419	2.97E-03	1.46E-01
Toe1	68276	0.545	0.260	6.24E-03	2.83E-01
Mlkl	74568	0.524	-0.047	8.17E-03	3.62E-01
Nasp	50927	0.523	0.275	8.25E-03	2.68E-01
Daxx	13163	0.517	0.809	8.83E-03	1.27E-02
Ifi204	15951	0.513	0.461	9.28E-03	1.18E-01
Aldoa	11674	0.458	0.891	1.77E-02	6.56E-03
Nat10	98956	0.457	0.850	1.80E-02	9.18E-03
Las1l	76130	0.443	1.088	2.11E-02	1.05E-03
Nol9	74035	0.396	0.860	3.43E-02	8.46E-03
Rpl8	26961	0.390	1.137	3.65E-02	6.28E-04
Set	56086	0.375	0.878	4.26E-02	7.30E-03
Anp32a	11737	0.364	0.849	4.72E-02	9.28E-03
Fhod1	234686	0.350	1.153	5.36E-02	5.31E-04
Rpl6	19988	0.339	0.850	5.94E-02	9.16E-03
Arbp	11837	0.324	0.945	6.80E-02	4.08E-03
Sart3	53890	0.309	0.885	7.71E-02	6.90E-03
Rpl7	19989	0.292	0.998	8.94E-02	2.50E-03
Pelp1	75273	0.243	1.337	1.31E-01	6.39E-05
Ranbp1	19385	0.017	1.334	4.64E-01	6.66E-05
Tex10	269536	-0.007	0.927	4.93E-01	4.79E-03
Gemin5	216766	-0.022	1.114	4.65E-01	8.00E-04
Ifi35	70110	-0.051	1.166	4.13E-01	4.61E-04
Cstf2	108062	-0.068	1.957	3.83E-01	5.89E-09
Nmi	64685	-0.071	0.967	3.79E-01	3.34E-03
Parp1	11545	-0.110	1.282	3.14E-01	1.25E-04
Cyb5r3	109754	-0.214	1.183	1.69E-01	3.84E-04
Arglu1	234023	-0.298	1.138	8.99E-02	6.23E-04
Lig3	16882	-0.354	0.908	5.51E-02	5.65E-03
Nf1	18015	-1.387	1.150	1.44E-10	5.51E-04
Cwf19l2	244672	-1.691	1.206	7.11E-15	2.98E-04

Supplementary Table 1.5a. DNA ligands.

Name	Sequences
ISD	5'-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3'
	5'-TGTAGATCATGTACAGATCAGTCATAGATCACTAGTAGATCTGTA-3'
HIV gag-100	5'-ATGGTTGTAGCTGTCCCAATATTTGTCTACAGCCTTCTGATGTCTCTAAAAGACCAGGATTAACTGCGAATCGTTCTAGCTCCCTGCTTACCCATACTA-3'
	5'-ATAGTATGGGTAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGTCTTTTAGAGACATCAGAAGGCTGTAGACAAATATTGGGACAGCTACAACCAT-3'
HSV60	5'-TAAGACACGATGCGATAAAATCTGTTTGTAAAATTTATTAAGGGTACAAATTGCCCTAGC-3'
	5'-GCTAGGGCAATTTGTACCCTTAATAAATTTTACAAACAGATTTTATCGCATCGTGTCTTA-3'
5'-biotin-ISD	5'-biotin-TACAGATCTACTAGTGATCTATGACTGATCTGTACATGATCTACA-3'
	5'-TGTAGATCATGTACAGATCAGTCATAGATCACTAGTAGATCTGTA-3'
5'-biotin-ISD-Ad1	5'-Biotin-TGAAGTCATGTGTGATATCACACTGCATGTGGCTAGACTTCATGG-3'
	5'-CCATGAAGTCTAGCCACATGCAGTGTGATATCACACATGACTTCA-3'
5'-biotin-ISD-act2	5'-Biotin-GTGGCTCCATCCTGGCCTCACTGTCCACCTTCCAGCAGATGTGGA-3'
	5'-TCCACATCTGCTGGAAGGTGGACAGTGAGGCCAGGATGGAGCCAC-3'
5'-biotin-gag-100	5'-biotin-ATGGTTGTAGCTGTCCCAATATTTGTCTACAGCCTTCTGATGTCTCTAAAAGACCAGGATTAACTGCGAATCGTTCTAGCTCCCTGCTTACCCATACTA-3'
	5'-ATAGTATGGGTAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGTCTTTTAGAGACATCAGAAGGCTGTAGACAAATATTGGGACAGCTACAACCAT-3'

Supplementary Table 1.5b. siRNA sequences.

Species	Gene	GeneID	Sense sequence	Notes
Mo	Abcf1	224742	Dharmacon siGENOME SMARTpool	si-0
Mo	Abcf1	224742	5'-AGGAAGUCCUGACUCGAAA-3'	si-1
Mo	Abcf1	224742	5'-CGATGATAGTGATGAGAGA-3'	si-2
Mo	IRF3	54131	5'-CGGACAAGCUUGUGAAGGA-3'	
Mo	Cdc37	12539	5'-GCGCCAAGCUGCGAAUAGA-3'	
Mo	Ptpn1	19246	5'-GACCACAGUCGGAUUAAAU-3'	
Mo	Tiparp	99929	5'-GAAACAUCACACCGUAUUG-3'	
Mo	Mdp1	67881	5'-AACGUUAACUCAAGGAUUA-3'	
Mo	Ppp6c	67857	5'-GCACGAAGGCUAUAAGUUU-3'	
Mo	Asb13	142688	5'-UAGAGAAAGUCGCCAAGUU-3'	
Mo	Trim56	384309	5'-GAAGCCAACUUGCGCUCUG-3'	
Mo	Usp49	224836	5'-GACCUGAAGUUGCUGAGAA-3'	
Mo	Reep4	72549	5'-GAUGAUCUGUCGCCUGGUA-3'	
Mo	Anp32a	11737	5'-CAATAGAGCCGCTGAAGAA-3'	
Mo	Cyb5r3	109754	5'-CCAAUGGGCUACUGGUCUA-3'	
Mo	Numa1	101706	5'-GGACGGCCAUUCUCUAGUA-3'	
Mo	Set	56086	5'-GGAAGATATTGATGAAGAA-3'	
Mo	Wdr77	70465	5'-GGGUGUCACUAGACUGGUA-3'	
Mo	Asf1a	66403	5'-AAUCUACAGUCCCUUCUUU-3'	
Mo	TBK1	56480	5'-CAGACUAGCUUAUAAUGAA-3'	
Mo	Mtmr3	74302	5'-GGGAAGAGGUGCCUGCUAU-3'	
Mo	Sting	72512	5'-GGAUCCGAAUGUUCAAUCA-3'	
Mo	RIG-I	230073	Dharmacon ON-TARGETplus SMARTpool	
Mo	Ifnar1	15975	Dharmacon siGENOME SMARTpool	
Mo	Jak1	16451	Dharmacon siGENOME SMARTpool	

Supplementary Table 1.5c. cDNA primers.

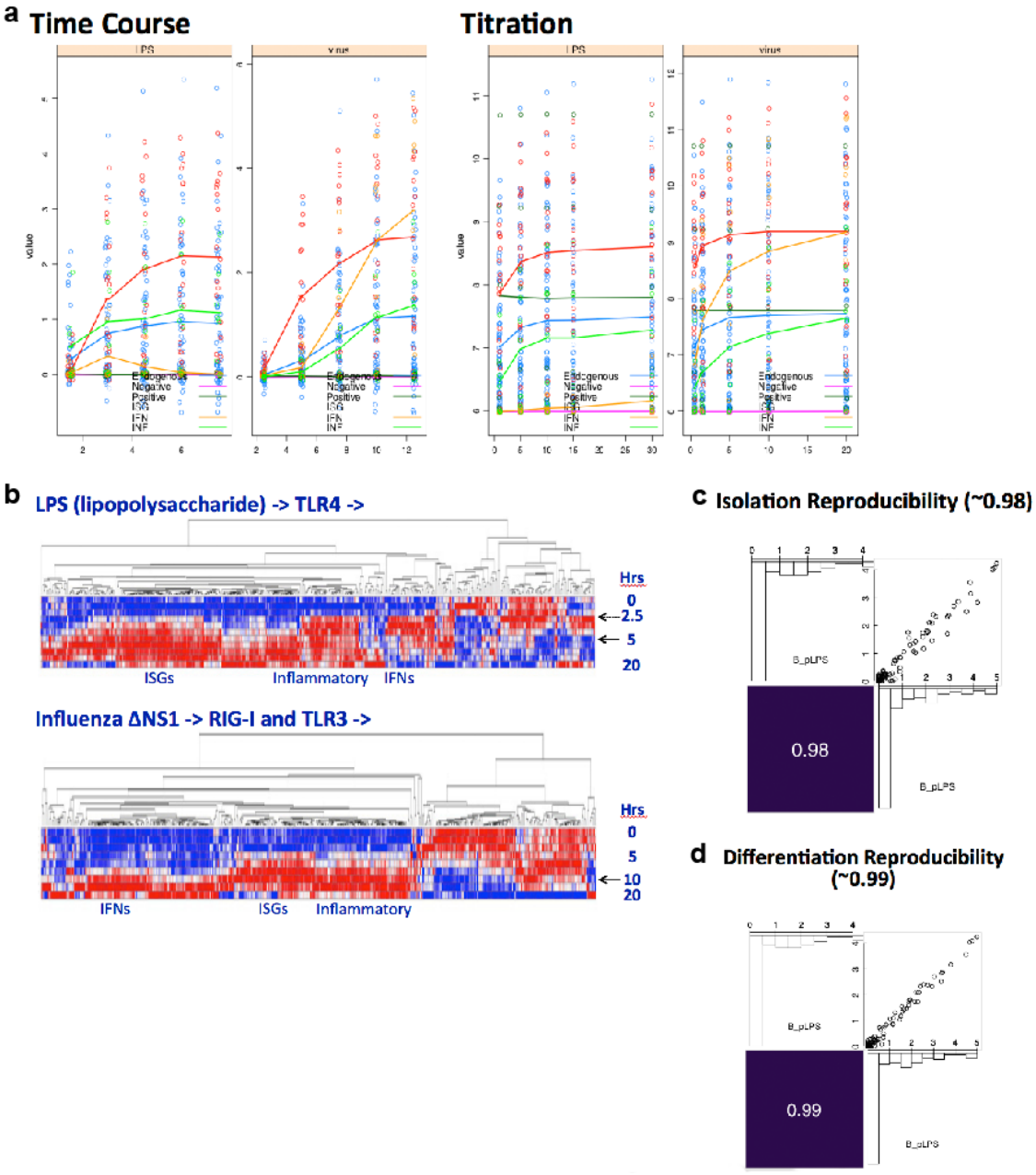
The following primers were used for cloning into pCW57d-P2AR:	
Abcf1 forward	5'-CATCATGCTAGCGCCGCCACCATGCCGAAGGGTCCCAAG-3'
Abcf1 reverse	5'-ATGATGACCGGTTCAATCCCGAGGACGGTTGAC-3'
Abcf1-HA reverse	5'-ATGATGACCGGTTCATGCGTAATCAGGCACATCGTAAGGGTATGCATCCCGAGGACGGTTGACCA-3'
Abcf1 rescue forward	5'-CAAAAGAGGTGTTAACACGTAAACAGCAGAAGTGCCGACG-3'
Abcf1 rescue reverse	5'-GCTGTTTACGTGTTAACACCTCTTTTGTTTGCTTTTCCGCTTGCTTG-3'
Irf3 forward	5'-CATCATGCTAGCGCCGCCACCATGGAAACCCCGAAACCGCGGA-3'
Irf3 reverse	5'-ATGATGACCGGTTCAGATATTTCCAGTGGCCTGGAAGT-3'
Irf3 rescue forward	5'-TCTGACTGATAAATTGGTCAAAGAGTACGTGGGGCAGGTGC-3'
Irf3 rescue reverse	5'-CACGTACTCTTTGACCAATTTATCAGTCAGAAACCCCTCAGGATCG-3'.
Renilla forward	5'-CATCATGCTAGCGCCGCCACCATGACTTCGAAAGTTTATGATCCAGAACAAAGGA-3'
Renilla-HA reverse	5'-ATGATGACCGGTTCATGCGTAATCAGGCACATCGTAAGGGTATGCTTGTTCATTTTTGAGAACTCGCTCAACGAAC-3'

The following primers were used for cloning into pLX301:	
Abcf1 forward	5'-CATCATGAATTCGCCGCCACCATGCCGAAGGGTCCCAAG-3'
Abcf1 reverse	5'-ATGATGCTCGAGTCAATCCCGAGGACGGTTGAC-3'
Abcf1-HA reverse	5'-ATGATGCTCGAGTCATGCGTAATCAGGCACATCGTAAGGGTATGCATCCCGAGGACGGTTGACCA-3'
Sting forward	5'-CATCATGAATTCgccGCCACCatgCCATACTCCAACCTGCATC-3'
Sting-HA reverse	5'-ATGATGCTCGAGtcaTGCGTAATCAGGCACATCGTAAGGGTATGCGATGAGGTCAGTGCGGA-3'

Supplementary Table 1.5d. qPCR primers.

Species	Gene	GeneID	Forward	Reverse
Mo	Gapdh	14433	5'-GGCAAATTCAACGGCACAGT-3'	5'-AGATGGTGATGGGCTTCCC-3'
Mo	Ifnb1	15977	5'-CTGGCTTCCATCATGAACAA-3'	5'-AGAGGGCTGTGGTGGAGAA-3'
Mo	Cxcl10	15945	5'-CCAAGTGCTGCCGTCATTTTC-3'	5'-GGCTCGCAGGGATGATTTCAA-3'
Mo	Abcf1	224742	5'-AGAAAGCCCGAGTTGTGTTTG-3'	5'-GCCCCCTTGTAGTCGTTGATG-3'
Mo	RIG-I	230073	5'-ACTTGGGTACAACATTGCGAG-3'	5'-GTTCACAAGAATCTGGGGTGTC-3'
Mo	Ifit1	15957	5'-CTGAGATGTCACTTCACATGGAA-3'	5'-GTGCATCCCCAATGGGTTCT-3'
Mo	Mx1	17857	5'-GACCATAGGGGTCTTGACCAA-3'	5'-AGACTTGCTCTTTCTGAAAAGCC-3'
Mo	Ifih1	71586	5'-AGATCAACACCTGTGGTAACACC-3'	5'-CTCTAGGGCCTCCACGAACA-3'
Mo	IL6	16193	5'-TAGTCCTTCCTACCCCAATTTCC-3'	5'-TTGGTCCTTAGCCACTCCTTC-3'
Mo	Irf7	54123	5'-GAGACTGGCTATTGGGGGAG-3'	5'-GACCGAAATGCTTCCAGGG-3'
Mo	Dhx58	80861	5'-GGAAGTGATCTTACCTGCTCTGG-3'	5'-TTGCCTCTGTCTACCGTCTCT-3'
Mo	Isg15	100038882	5'-GGTGTCCGTGACTAACTCCAT-3'	5'-TGGAAAGGGTAAGACCGTCCT-3'
Mo	Ifi44	99899	5'-AACTGACTGCTCGCAATAATGT-3'	5'-GTAACACAGCAATGCCTCTTGT-3'
Mo	Stat1	20846	5'-TCACAGTGGTTCGAGCTTCAG-3'	5'-GCAAACGAGACATCATAGGCA-3'
Mo	Ccl5	20304	5'-GCTGCTTTGCCTACCTCTCC-3'	5'-TCGAGTGACAAACACGACTGC-3'
Mo	Ptpn1	19246	5'-GGAACTGGGCGGCTATTTACC-3'	5'-CAAAAGGGCTGACATCTCGGT-3'
Mo	Tiparp	99929	5'-GCCAGACTGTGTAGTACAGCC-3'	5'-GGGTTCCAGTTCCCAATCTTTT-3'
Mo	Ppp6c	67857	5'-CCGCTGGATCTGGACAAGTAT-3'	5'-ACACTGGCTGAACATTCGACT-3'
Mo	Mdp1	67881	5'-TTGATCTGGATTACACGCTCTGG-3'	5'-CCATCGCTGCTCTTGTGGAAT-3'
Mo	Asb13	142688	5'-TCCTGGGAGATGTGGGTTTCT-3'	5'-AAGGGGTGTGATCGAGTCCA-3'
Mo	Trim56	384309	5'-AAGACTCCTCCCCAACTCTG-3'	5'-GGCAATAGGTATGTAGGCATGG-3'
Mo	Usp49	224836	5'-AGTTCCGGGAATGTTTCCTGA-3'	5'-CTCCTTACTGACAACTCTGCG-3'
Mo	Reep4	72549	5'-GCCTGGTAGTGCTCATATTTGG-3'	5'-GCCATGAAGATCGCAAAGACAA-3'
Mo	Cyb5r3	109754	5'-CAGGGCTTCGTGAATGAGGAG-3'	5'-TCCACACATCAGTATCAGCGG-3'
Mo	Numa1	101706	5'-CCCAAGGGAGGAATAGCTTCT-3'	5'-CTCTGCGATGCGGTTCCAA-3'
Mo	Wdr77	70465	5'-CTTGCTGTGCTGGATTCAAGC-3'	5'-CAACTGTGGTAAGAAGGGAGTG-3'
Mo	Irf3	54131	5'-GAGAGCCGAACGAGGTTCAG-3'	5'-CTTCCAGGTTGACACGTCCG-3'
Mo	Asf1a	66403	5'-GTGGTGCTGGATAACCCGTC-3'	5'-GGGACCCACTAAAACAGAGTCTA-3'
Mo	Tbk1	56480	5'-ACTGGTGATCTCTATGCTGTCA-3'	5'-TTCTGGAAGTCCATACGCATTG-3'
Mo	Mtmr3	74302	5'-ATGACTCGTTGGCTACCTGAC-3'	5'-GAACCGGAACCTTCTGGTTAC-3'
Mo	Sting	72512	5'-GGTCACCGCTCCAAATATGTAG-3'	5'-CAGTAGTCCAAGTTCGTGCGA-3'
Mo	Cdc37	12539	5'-GACTACAGCGTTTGGGATCAC-3'	5'-CCCCGGTCCAGTTCCTCTT-3'

Appendix B: Supporting information for Chapter 2



Supplementary Figure 2.1. A robust system to quantitatively assess innate immune responses in humans. (**a**) Time course and titration curves of MoDCs pooled from 5 donors and stimulated with LPS or infected with influenza virus. Lysates were run on a pilot Nanostring codeset. (**b**) Time course of MoDCs pooled from 13 donors and stimulated with LPS or infected with influenza virus. Global gene-expression was measured by microarray, and heatmap of selected genes is shown. (**c**) To estimate the technical reproducibility of the assay, PBMCs were divided into two samples and the assay was run in parallel. Monocytes from the two samples were isolated, differentiated into MoDCs, and stimulated with LPS; lysates were then run on Nanostring. $R^2 = 0.98$. (**d**) To estimate the technical reproducibility of steps following monocyte isolation, monocytes were divided into two samples and the assay was run in parallel as above. $R^2 = 0.99$.